

Portafolio de Evidencias

Extracción de Conocimiento en Bases de Datos

Carrera: Ingeniería en Desarrollo y Gestión de software

Alumno: Daniel Contreras Zamarripa

Profesor: MGTI. María Eugenia Guerrero Chan

INDICE GENERAL

Contenido

Unidad 1 Introducción al análisis de datos	3
Resumen.....	3
practica de ejercicios.....	8
Ejercicios de frecuencias	17
Examen	20

Unidad 1 Introducción al análisis de datos

Resumen

Instrumento:	Resumen
---------------------	---------

Alumno: Daniel Contreras Zamarripa	Fecha: 02/05/2024
Carrera: Ingeniería en Desarrollo y Gestión de Software	Grupo: IDGS91
Asignatura: Extracción del Conocimiento en Base de Datos	Unidad temática: I. Introducción al análisis de datos
Profesor: MGTI María Eugenia Guerrero Chan	

Instrucciones:

- Escribir una *Introducción* sobre las Metodologías para el análisis de datos

I. Título: Metodologías para el análisis de datos

II. Contenido (Introducción y Desarrollo):

Las metodologías para el análisis de datos son fundamentales en la era de la información, donde el volumen, la variedad y la velocidad de los datos generados superan cualquier expectativa previa. Estas metodologías proporcionan un marco estructurado para recolectar, procesar y analizar datos, transformándolos en información valiosa para la toma de decisiones.

El análisis de datos comienza con la definición clara de objetivos, seguido por la selección de las herramientas y técnicas adecuadas. Métodos como la estadística descriptiva y la inferencial permiten resumir y explorar los datos, así como hacer predicciones y generalizaciones. Por otro lado, la minería de datos utiliza algoritmos para descubrir patrones y relaciones ocultas.

La visualización de datos es otra piedra angular, que facilita la comprensión y comunicación de los resultados a través de gráficos y tableros interactivos. Además, metodologías como el Aprendizaje Automático y la Inteligencia Artificial están revolucionando el campo, permitiendo análisis predictivos y prescriptivos más sofisticados.

Finalmente, la gestión y el mantenimiento de la calidad de los datos son críticos para garantizar la precisión y la confiabilidad del análisis. Esto incluye la limpieza de datos, la detección de anomalías y la integración de diversas fuentes de datos.

1. Que es un Proyecto de Análisis de Datos

Un Proyecto de Análisis de Datos es una iniciativa estructurada que busca extraer información significativa de conjuntos de datos complejos para apoyar la toma de decisiones, identificar tendencias, y resolver problemas específicos. Este tipo de proyecto implica varias etapas, que incluyen la recolección, limpieza, exploración, análisis e interpretación de los datos.

En la práctica, un proyecto de análisis de datos puede abordar desde simples consultas estadísticas hasta modelos predictivos avanzados utilizando técnicas de Aprendizaje Automático. Los proyectos pueden variar en complejidad y ser aplicados en distintos sectores, como negocios, ciencia, tecnología, salud y más.

2. Características de un Proyecto de Análisis de Datos.

- **Objetivo Claro:** Todo proyecto de análisis de datos debe comenzar con una definición precisa de los objetivos. Esto incluye qué preguntas se buscan responder y qué problemas se pretenden resolver
- **Recolección de Datos:** Se deben recopilar datos relevantes y de calidad. Esto puede implicar el uso de encuestas, bases de datos existentes, o la generación de nuevos datos a través de la experimentación.

- Limpieza y Preparación: Los datos recogidos suelen requerir limpieza para eliminar errores o inconsistencias, y preparación para convertirlos en un formato adecuado para el análisis.
- Análisis: Utilización de técnicas estadísticas, algoritmos de minería de datos, o modelos de inteligencia artificial para examinar los datos y extraer patrones, tendencias y relaciones

3. Que es una Metodología para el análisis de datos.

Una Metodología para el análisis de datos es un conjunto de procedimientos y técnicas estructuradas que guían la recolección, procesamiento y análisis de datos para descubrir información útil, extraer conocimientos y tomar decisiones informadas. Incluye pasos sistemáticos que van desde la definición de objetivos hasta la interpretación y comunicación de los resultados.

Las metodologías pueden variar según el contexto y las necesidades específicas, pero generalmente incluyen:

- Recopilación de Requisitos: Definir lo que se necesita saber y por qué es importante.
- Recolección de Datos: Obtener datos de fuentes confiables y pertinentes.
- Procesamiento de Datos: Convertir los datos crudos en un formato que pueda ser analizado.
- Limpieza de Datos: Eliminar inexactitudes o datos irrelevantes para asegurar la calidad.
- Análisis/Comunicación: Aplicar técnicas estadísticas y de visualización para interpretar los datos y comunicar los hallazgos.

4. Principales Metodologías o Métodos de análisis de datos (definición de cada uno de ellos y ejemplos donde se aplican).

Las principales metodologías de análisis de datos son técnicas que permiten a los analistas extraer insights y conocimientos a partir de datos brutos. Aquí hay una descripción de cada una y ejemplos de su aplicación.

- **Análisis Descriptivo:** Proporciona una visión general de lo que ha ocurrido en la historia de los datos recogidos. Utiliza estadísticas para describir las características principales de un conjunto de datos, como la media, mediana, moda y desviación estándar. Ejemplo: En el ámbito empresarial, se utiliza para resumir las ventas anuales y entender las tendencias generales.
- **Análisis Predictivo:** Utiliza modelos estadísticos y de aprendizaje automático para hacer predicciones sobre eventos futuros basándose en datos históricos. Ejemplo: En la industria financiera, se aplica para predecir la dirección de los precios de las acciones o la probabilidad de incumplimiento de un crédito.
- **Análisis Exploratorio:** Se centra en descubrir patrones, anomalías, relaciones o tendencias desconocidas en los datos. Ejemplo: En la investigación científica, se usa para explorar grandes conjuntos de datos genómicos en busca de posibles asociaciones con enfermedades.
- **Análisis Prescriptivo:** Va más allá de predecir lo que podría suceder, sugiriendo acciones específicas que podrían influir en los resultados. Ejemplo: En la logística, puede ayudar a determinar las rutas de entrega más eficientes para reducir costos y tiempos.
- **Análisis de Redes:** Examina las relaciones entre entidades en una red. Ejemplo: En las redes sociales, se utiliza para identificar influenciadores clave y la estructura de las comunidades.
- **Análisis Espacial:** Se refiere al examen de la ubicación geográfica y las características de los fenómenos naturales o humanos. Ejemplo: En la planificación urbana, se usa para analizar patrones de crecimiento de la ciudad y planificar infraestructuras.

III. Bibliografía:

References

Aguilera, R. (2023, September 14). 25 proyectos para analista de datos. *Desafío Latam - Academia de Talentos Digitales*.

<https://blog.desafiolatam.com/25-proyectos-para-analista-de-datos/>

Análisis de datos: Concepto, metodología y técnicas. (2023, July 18).

InnovaciónDigital360. <https://www.innovaciondigital360.com/big-data/analisis-de-datos-tecnicas-y-metodologias-para-la-aplicacion-de-analytics/>

Deep Learning y Innovación Digital. (n.d.). InnovaciónDigital360. Retrieved May 2, 2024, from <https://www.innovaciondigital360.com/i-a/deep-learning/>

Rodríguez, J. (2024, January 17). ✓ *Qué es el análisis de datos: definición, etapas y ejemplos*. Mundo Posgrado. <https://www.mundoposgrado.com/que-es-el-analisis-de-datos/>

(N.d.). Datacamp.com. Retrieved May 2, 2024, from

<https://www.datacamp.com/es/blog/data-analytics-projects-all-levels>

practica de ejercicios

Instrumento	<i>Práctica de ejercicios</i>
--------------------	-------------------------------

Alumno: Daniel Contreras Zamarripa	Fecha: 23/05/2024
Carrera: Ingeniería en Desarrollo y Gestión de Software	Grupo: IDGS91
Asignatura: Extracción de Conocimiento en Bases de Datos.	Unidad temática: I. Introducción al análisis de datos
Profesor: MGTI. María Eugenia Guerrero Chan	

I.- Ejercicios a resolver:

Instrucciones:





Observa la tabla siguiente e investiga la información que se pide y con base a la investigación llena la tabla.

VALOR (50%)

Tabla 1.- Llena la siguiente tabla con la información que se pide.

	Características	Casos de aplicación	Lenguajes y herramientas
Inteligencia Artificial	<ul style="list-style-type: none"> Imita el modo de pensar Humano. Automatiza Procesos. Gestiona una enorme cantidad de datos. 	<p>Amazon: La compañía de comercio electrónico utiliza la inteligencia artificial para hacer recomendaciones personalizadas a sus clientes y para predecir la demanda de productos, lo que les permite planificar y optimizar su inventario.</p> <p>BMW: El fabricante de automóviles utiliza la inteligencia artificial para optimizar su proceso de producción, reducir costos y mejorar la calidad de sus vehículos.</p>	<p>DALL-E: herramientas para crear imágenes. Con la capacidad de generar imágenes a partir de descripciones textuales.</p> <p>Fireflies: creada para grabar y transcribir conversaciones, es una herramienta muy útil para las conversaciones en tiempo real. Además, su asistente de inteligencia artificial transforma el diálogo en datos de valor, facilitando transcripciones ágiles y precisas de cada conversación.</p>
Big Data	<ul style="list-style-type: none"> Veracidad: la información que abarca el big data debe ser verdadera y comprobable. Variedad: la información que se obtiene de distintas fuentes, lo que permite que no haya sesgo, pero tampoco 	<p>Desarrollo de productos: Empresas como Netflix y Procter & Gamble usan big data para prever la demanda de los clientes. Construyen modelos predictivos para nuevos productos y servicios clasificando atributos clave de productos anteriores y actuales, y modelando la relación entre dichos atributos y el éxito comercial.</p> <p>Mantenimiento predictivo: Analizan datos estructurados y no estructurados, como registros, datos de sensores, mensajes de error y temperaturas de motor, para identificar indicadores de fallos mecánicos potenciales antes de que ocurran.</p>	<p>Hadoop: La biblioteca de software es un marco de big data. Permite el procesamiento distribuido de grandes conjuntos de datos en clusters de computadoras. Es una de las mejores herramientas de big data diseñada para escalar desde servidores únicos hasta miles de máquinas. Aportando flexibilidad en el procesamiento de datos.</p> <p>Pentaho es una suite de software de código abierto utilizada para el análisis de datos, la integración de datos y la generación de informes en el ámbito del Business Intelligence (BI). Proporciona una plataforma completa que abarca desde la extracción y</p>

	<p>límites de lo que se aprende del análisis.</p> <p>+ Valor: Al final de su análisis, el big data ayudará a tomar acciones y a reconocer oportunidades para mejorar la empresa en varios de sus niveles.</p>		<p>transformación de datos hasta la visualización y el análisis de datos.</p>
Machine Learning	<p>+ Capacidad de procesar grandes cantidades de datos.</p> <p>+ Aprendizaje autónomo.</p> <p>+ Automatización de tareas.</p> <p>+ Adaptabilidad a diferentes situaciones.</p>	<p>Recomendaciones personalizadas que fidelizan a los clientes: Como ejemplos tenemos las sugerencias o recomendaciones de compra en plataformas de comercio electrónico. En un principio se analiza el historial de interacciones de cada usuario (compras, visualizaciones de producto, búsquedas, etc.) y se compara con otros usuarios parecidos.</p> <p>Detección de operaciones fraudulentas: la detección de fraude es una preocupación fundamental del proceso. Se están capacitando modelos predictivos para evaluar los riesgos mediante el análisis de casos de fraude financiero. Para determinar los riesgos en función de las</p>	<p>Mangahigh: El proyecto se enfoca en el área de las matemáticas. El docente puede personalizar las instrucciones en un proceso de tres pasos: Asignar actividades, revisión de rendimiento y diferenciación de instrucciones.</p> <p>Gracias a la información y datos que provee Mangahigh el docente puede analizar los resultados del grupo de estudiantes y cambiar instrucciones para mejorar el proceso de aprendizaje.</p> <p>Scikit-Learn: proporciona una amplia gama de herramientas para el aprendizaje automático, como algoritmos de clasificación, regresión, agrupamiento y</p>

		<p>características de la aplicación, los modelos luego analizan datos como el proceso de solicitud del crédito y la probabilidad del fraude.</p>	<p>reducción de dimensionalidad. También incluye una variedad de herramientas de preprocesamiento para la normalización, escalado y codificación de datos. TensorFlow: permite construir y entrenar redes neuronales para detectar patrones y razonamientos usados por los humanos. Funciona en diversas plataformas y tiene aplicaciones en campos como el reconocimiento de imágenes y el análisis de sentimientos.</p>
Data Mining	<ul style="list-style-type: none">  Análisis de Grandes Volúmenes de Datos.  Extracción de Patrones y Conocimiento.  Predicción y Clasificación.  Visualización de Datos 	<p>Inteligencia artificial</p> <p>Un sistema se vuelve artificialmente inteligente alimentándolo con patrones relevantes. Estos patrones provienen de los resultados de la data mining. Los resultados de los sistemas artificialmente inteligentes también se analizan para determinar su relevancia utilizando técnicas de data mining.</p> <p>previsión de tendencias: Al analizar grandes conjuntos de datos, la minería de datos puede identificar patrones y tendencias que pueden no ser evidentes a simple vista. Estos patrones pueden luego ser utilizados para hacer predicciones sobre futuras tendencias.</p>	<p>Rapid Miner: En esta herramienta de data mining se pueden tanto minar datos como realizar análisis predictivos de forma gratuita, está ganando bastante cuota de mercado estos últimos años debido a la facilidad para crear chatbots, aunque se puede utilizar para innumerables proyectos.</p> <p>R es un lenguaje estadístico que nació específicamente para el análisis de datos. Los que tienen perfiles técnicos o de negocios suelen preferir R por delante de Python. Esto es porque sus gráficas para la visualización de datos son muy intuitivas y se pueden crear informes y visualizaciones de datos de gran calidad.</p>

PRÁCTICA DE EJERCICIOS

Sistema de Gestión de la Calidad

Fecha de emisión:

10/01/2013

Revisión: 01

Página 12 de 20

			<p>Python: este lenguaje de programación no es tan específico en la parte estadística como R, pero ha ido introduciéndose en el mundo de data science con el tiempo. Normalmente, las personas que vienen de perfiles matemáticos o del mundo del desarrollo escogen esta herramienta de minería de datos antes que R.</p>
--	--	--	--



Con base a la información que pusiste en la tabla **selecciona** y **justifica** con tus propias palabras una herramienta de Inteligencia Artificial, Machine Learning, Data Mining y una de Big Data como apoyo para la toma de decisiones.

La justificación es de manera individual, es decir; una para Inteligencia Artificial, Machine Learning, Data Mining y una de Big Data.

II.-Procedimientos y resultados:

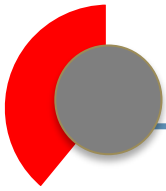
Inteligencia Artificial

DALL-E, una herramienta de inteligencia artificial para la generación de imágenes, es increíblemente útil por varias razones, destacándose por su capacidad para transformar descripciones textuales detalladas en imágenes realistas y artísticas. Esta capacidad ofrece ventajas significativas en diversos campos, desde el diseño gráfico hasta la educación y la investigación científica.

Primero, en el ámbito del diseño gráfico y la publicidad, DALL-E facilita la creación de imágenes únicas y personalizadas sin necesidad de contratar ilustradores o dedicar horas al dibujo manual. Los diseñadores pueden simplemente describir la imagen que desean y dejar que DALL-E genere una variedad de opciones. Esta rapidez y eficiencia no solo ahorra tiempo y recursos, sino que también permite iterar rápidamente sobre conceptos y obtener retroalimentación inmediata, mejorando el proceso creativo y la calidad del producto final.

Además, en la educación, DALL-E puede ser una herramienta poderosa para ilustrar conceptos complejos que pueden ser difíciles de explicar solo con palabras. Por ejemplo, en las ciencias, los maestros pueden generar visualizaciones detalladas de procesos biológicos, químicos o físicos que ayuden a los estudiantes a comprender mejor los temas. Esto es particularmente útil en el aprendizaje visual, donde las imágenes pueden ayudar a reforzar la comprensión y retención de la información.

No obstante, DALL-E también tiene aplicaciones significativas en la creación de contenido digital y multimedia. Los creadores de contenido pueden usar esta herramienta para generar imágenes personalizadas que se alineen perfectamente con su narrativa o mensaje, mejorando así la calidad y el atractivo visual de su contenido. Esta personalización puede hacer que el contenido sea más atractivo y relevante para el público objetivo.



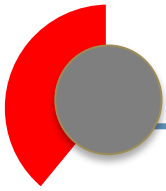
Big Data

Hadoop permite el almacenamiento escalable y económico de datos. Con su sistema de archivos distribuido (HDFS), las organizaciones pueden almacenar petabytes de datos de manera eficiente en una red de servidores estándar, reduciendo así los costos en comparación con los sistemas de almacenamiento tradicionales. Esta capacidad es crucial para las empresas que generan y deben gestionar enormes cantidades de datos provenientes de diversas fuentes como redes sociales, dispositivos IoT y transacciones comerciales.

Además, Hadoop proporciona un marco robusto para el procesamiento de datos en paralelo. Utilizando el modelo de programación MapReduce, Hadoop divide grandes tareas de procesamiento en tareas más pequeñas que se ejecutan simultáneamente en diferentes nodos de la red. Esto no solo acelera el tiempo de procesamiento de datos masivos, sino que también mejora la eficiencia y la capacidad de manejar picos de carga, lo que es vital para análisis en tiempo real y toma de decisiones rápidas.

Enfocándonos en términos de flexibilidad, Hadoop se destaca por su capacidad para procesar y analizar datos estructurados y no estructurados. Esta versatilidad permite a las organizaciones integrar y analizar datos de diversas fuentes y formatos, desde bases de datos tradicionales hasta datos no estructurados como correos electrónicos, videos y registros de sensores. Esto es particularmente útil en sectores como la salud, finanzas y telecomunicaciones, donde la variedad de datos es amplia y compleja.

Hadoop también promueve la innovación y el desarrollo continuo gracias a su ecosistema abierto y en crecimiento. Con componentes adicionales como Hive, Pig, HBase y Spark, los usuarios pueden realizar consultas SQL, procesamiento en tiempo real, almacenamiento de datos en tiempo real y análisis de flujos de datos, entre otras capacidades. Este ecosistema permite a las organizaciones adaptarse rápidamente a nuevas tecnologías y métodos analíticos.



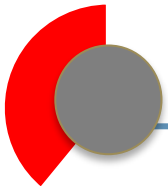
Machine Learning

Scikit-Learn destaca por su facilidad de uso y accesibilidad. La biblioteca ofrece una interfaz simple y coherente para una amplia variedad de algoritmos de machine learning, lo que permite a los usuarios, incluso aquellos con poca experiencia en programación o machine learning, construir y aplicar modelos con relativa facilidad. La documentación exhaustiva y los numerosos ejemplos proporcionados hacen que la curva de aprendizaje sea más suave, permitiendo a los desarrolladores concentrarse en la solución de problemas específicos en lugar de en la implementación de algoritmos desde cero.

Además, Scikit-Learn incluye una amplia gama de algoritmos para tareas de clasificación, regresión, clustering y reducción de dimensionalidad. Esto proporciona a los usuarios la flexibilidad de elegir el método más adecuado para sus necesidades específicas sin tener que recurrir a múltiples bibliotecas. La consistencia en la API de Scikit-Learn también facilita la comparación y evaluación de diferentes modelos, permitiendo un enfoque iterativo y experimental en la construcción de soluciones de machine learning.

Scikit-Learn también proporciona herramientas robustas para la validación cruzada y la optimización de hiperparámetros, lo que es crucial para el desarrollo de modelos de alto rendimiento. La validación cruzada ayuda a evaluar la eficacia del modelo de manera más precisa, reduciendo el riesgo de sobreajuste. La optimización de hiperparámetros, mediante técnicas como Grid Search y Randomized Search, permite encontrar las mejores configuraciones de parámetros para maximizar el rendimiento del modelo.

En el ámbito empresarial, Scikit-Learn es altamente valorado por su eficiencia y escalabilidad. Aunque está diseñado principalmente para entornos de investigación y desarrollo, su implementación ligera y eficiente permite el despliegue de modelos en producción con facilidad. Las empresas pueden integrar modelos de machine learning en sus sistemas existentes para automatizar procesos, mejorar la toma de decisiones y generar insights valiosos a partir de grandes volúmenes de datos.



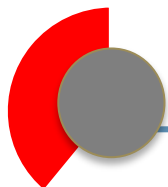
Data Mining

RapidMiner se distingue por su interfaz de usuario intuitiva y amigable. A diferencia de otras herramientas de data mining que pueden requerir un conocimiento profundo de programación, RapidMiner permite a los usuarios diseñar y ejecutar procesos de análisis de datos mediante una interfaz gráfica de arrastrar y soltar. Esta característica democratiza el acceso al análisis avanzado de datos, permitiendo a los analistas y científicos de datos centrarse en la interpretación de los resultados en lugar de en la codificación detallada.

Además, ofrece una extensa gama de funciones y algoritmos para data mining. La plataforma incluye más de 1500 operadores para tareas como la preparación de datos, modelado, evaluación y despliegue. Esto permite a los usuarios abordar una variedad de problemas de análisis de datos, desde la clasificación y regresión hasta el clustering y la reducción de dimensionalidad, utilizando una única herramienta integrada. La amplitud de funcionalidades y la capacidad de integrar diversos algoritmos facilita la creación de flujos de trabajo complejos y personalizados que se ajustan a las necesidades específicas de cada proyecto.

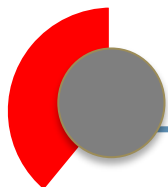
Otro aspecto clave de RapidMiner es su capacidad para manejar grandes volúmenes de datos y su integración con múltiples fuentes de datos. La plataforma soporta la conexión a bases de datos tradicionales, big data y servicios en la nube, lo que permite a los usuarios importar y exportar datos de manera eficiente. Esta conectividad y flexibilidad en el manejo de datos son cruciales en entornos empresariales donde los datos provienen de diversas fuentes y formatos.

En términos de colaboración y escalabilidad, RapidMiner proporciona funcionalidades avanzadas que facilitan el trabajo en equipo y el despliegue de modelos en producción. Los equipos pueden compartir y reutilizar flujos de trabajo, lo que mejora la colaboración y la eficiencia. Además, la plataforma soporta la implementación de modelos en entornos productivos, asegurando que los análisis y predicciones puedan ser integrados fácilmente en los sistemas operativos de la empresa.

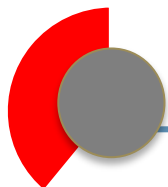


Ejercicios de frecuencias

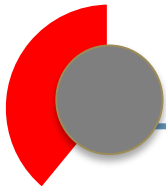
Autos Vendidos	Cantidad	frecuencia absoluta (fi)	frecuencia acumulada (Fi)	frecuencias relativas (ni)	frecuencia relativa acumulada (Ni)
0	1	7	7	0.233	0.233
0	2	7	14	0.233	0.467
0	3	5	19	0.167	0.633
0	4	3	22	0.100	0.733
0					
0					
0		Cantidad Total	30		
0		Numero Menor	0		
1		Numero Mayor	4		
1					
1					
1					
1					
1					
1					
2					
2					
2					
2					
2					
2					
2					
3					
3					
3					
3					
3					
4					
4					
4					



Calificaciones	Calificacion	frecuencia absoluta (fi)	frecuencia acumulada (Fi)	frecuencias relativas (ni)	frecuencia relativa acumulada (Ni)
9	6	1	1	0.043	0.043
6	7	5	6	0.217	0.261
7	8	6	12	0.261	0.522
7	9	7	19	0.304	0.826
7	10	4	23	0.174	1.000
7					
7		Cantidad Total	23		
8		Calificacion Mer	6		
8		Calificacion May	10		
8					
8					
8					
8					
9					
9					
9					
9					
9					
9					
10					
10					
10					
10					



Peso Perdido	Peso	frecuencia absoluta (fi)	frecuencia acumulada (Fi)	frecuencias relativas (ni)	frecuencia relativa acumulada (Ni)
0.2	0.2	1	1	0.063	0.063
3.5	3.5	1	2	0.063	0.125
4.3	4.3	1	3	0.063	0.188
4.6	4.6	1	4	0.063	0.250
6.4	6.4	1	5	0.063	0.313
6.5	6.5	1	6	0.063	0.375
8.4	8.4	1	7	0.063	0.438
8.7	8.7	1	8	0.063	0.500
9.1	9.1	1	9	0.063	0.563
9.6	9.6	1	10	0.063	0.625
12.1	12.1	2	12	0.125	0.750
12.1	14.3	1	13	0.063	0.813
14.3	15.2	1	14	0.063	0.875
15.2	16.1	1	15	0.063	0.938
16.1	19.8	1	16	0.063	1.000
19.8					
Cantidad Total	16				
Calificacion Menor	0.2				
Calificacion Mayor	19.8				



Examen

Instrumento	Examen
Alumno: <i>Donal Contreras Zamora</i>	Fecha: <i>28/05/2024</i>
Carrera: Ingeniería en Desarrollo y Gestión de Software	Grupo:
Asignatura: Extracción de Conocimiento en Bases de Datos.	Unidad temática: I. Introducción al análisis de datos
Profesor: MGTI. María Eugenia Guerrero Chan	

I. Reactivos

1. Hablar de inteligencia artificial (IA) es algo tan sencillo como hablar de:

- a) Big data
- b) Variedad
- c) ☒ Máquinas inteligentes
- d) Disponibilidad de grandes volúmenes de datos

2. Se refiere a los datos que son tan grandes, rápidos o complejos que es difícil o imposible procesarlos con los métodos tradicionales.

- a) Inteligencia artificial
- b) ☒ Big data
- c) Machine learning
- d) Data mining

3. Es un campo de la estadística y las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos.

- a) Big data
- b) Machine learning
- c) ☒ Data mining
- d) Inteligencia artificial

4. Es una rama de la inteligencia artificial basada en la idea de que los sistemas pueden aprender de datos, identificar patrones y tomar decisiones con mínima intervención humana.

- a) Data mining
- b) ☒ Machine learning
- c) Big Data
- d) Inteligencia artificial