1. **DATOS GENERALES**

|  |  |  |
| --- | --- | --- |
| **RESPONSABLES EQUIPO DE PRODUCCIÓN** |  | Registro revisión |
| **Líder de producción** | Olga Constanza Bermúdez Jaimes | |
| **Nombre Asignatura:** | Big Data y Analítica de Datos | |
| **Experto Disciplinar** | Iván Guillermo Ortega Diez | Diciembre 2024 |
| **Asesor Tecno pedagógico** | Elizabeth Bermúdez Díez | Diciembre 2024 |
| **Diseñador Instruccional** | Paola Alexandra Moya Peralta | Diciembre 2024 |
| **Validador de contenidos** | María Sus | Diciembre 2024 |
| **Diseñador gráfico** |  |  |
| **Maquetador web** |  |  |

|  |  |
| --- | --- |
| **U1** | Fundamentos de analítica de datos |
| **Resumen banner** | La analítica de datos permite identificar tendencias y oportunidades, mediante la recolección y análisis de información. Se divide en analítica descriptiva, predictiva y prescriptiva. La descriptiva, analiza datos históricos para comprender qué ocurrió, usando herramientas como Power BI y K-Means. La predictiva, anticipa eventos futuros mediante *machine learning* y modelos estadísticos. La minería de datos, procesa información estructurada y no estructurada, combinando métricas cuantitativas y cualitativas, con herramientas como Hadoop y Spark. Sus aplicaciones incluyen optimización de recursos, detección de fraudes y toma de decisiones estratégicas.  *Dominar la analítica de datos, es el primer paso para transformar desafíos en oportunidades y construir un futuro guiado por decisiones inteligentes y estratégicas.* |

1. **INTRODUCCIÓN UNIDAD 2**

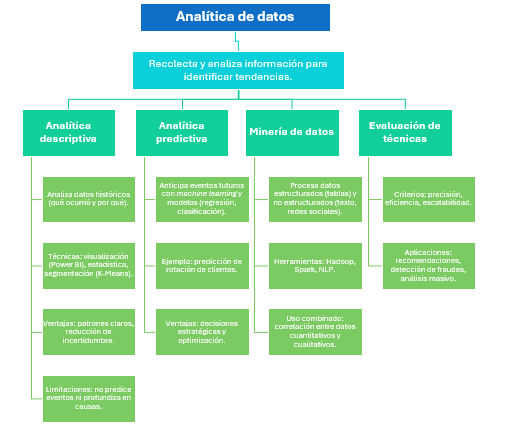
La analítica de datos, es fundamental para transformar grandes volúmenes de información en conocimiento útil que permita a las organizaciones tomar decisiones informadas. Desde el análisis descriptivo, que sintetiza datos históricos para identificar patrones y tendencias, hasta el predictivo, que anticipa eventos futuros mediante modelos avanzados, cada enfoque contribuye a optimizar estrategias y procesos claves. Estas técnicas ayudan a comprender el pasado y proyectar escenarios futuros, permitiendo a las organizaciones adaptarse mejor a los cambios del entorno.

|  |  |
| --- | --- |
| https://www.freepik.es/vector-gratis/composicion-isometrica-programacion-cubos\_23246946.htm#fromView=search&page=1&position=2&uuid=fee42ad3-a444-4f2d-8a3c-0b8c0ac9868f | Además, la minería de datos amplía las posibilidades al trabajar con datos estructurados y no estructurados, ofreciendo un análisis integral que combina métricas cuantitativas y percepciones cualitativas. Con el apoyo de herramientas tecnológicas como Power BI, Tableau o Apache Spark, es posible extraer *insights* accionables de información compleja, mejorando la eficiencia operativa y la capacidad de respuesta, frente a desafíos específicos. |

**∫**

1. **SÍNTESIS UNIDAD 2**

La unidad 2 se centra en los principios esenciales de la analítica de datos, abordando las técnicas y herramientas claves para transformar grandes volúmenes de datos estructurados y no estructurados, en información valiosa. Este enfoque proporciona a los estudiantes las competencias necesarias para aplicar métodos analíticos, en diversos contextos, promoviendo la toma de decisiones fundamentadas en evidencia.



**Desarrollo contenido**

**1. Tipos de analítica**

La analítica de datos es un conjunto de acciones que permiten recolectar y aprovechar información para identificar tendencias y oportunidades. A continuación, se presenta una descripción detallada de la analítica descriptiva, combinando párrafos explicativos, características principales y un cuadro resumen.

**1.1. Analítica descriptiva**

La analítica descriptiva es el punto de partida en el análisis de datos. Su propósito es sintetizar y visualizar información histórica para comprender qué ocurrió y por qué, estableciendo una base sólida para decisiones estratégicas y futuras proyecciones. Este enfoque analiza eventos pasados mediante la identificación de patrones, tendencias y relaciones, transformando datos complejos en información clara y accionable.

**Características principales de la analítica descriptiva**

La analítica descriptiva es una herramienta esencial para comprender y resumir eventos pasados mediante el análisis de datos históricos:

|  |  |  |
| --- | --- | --- |
| **SLIDE** | | |
| **Descripción de eventos pasados** | Permite resumir e interpretar datos históricos. Por ejemplo, al analizar el comportamiento de ventas mensuales para identificar picos en temporadas específicas como Navidad. | <https://www.freepik.es/vector-gratis/infografia-puesta-marcha-compania_846253.htm#fromView=search&page=1&position=1&uuid=cd9ea0ed-2844-48ad-b1df-f392514d53ac> |
| **Enfoque en patrones** | Analiza distribuciones, tendencias y correlaciones entre variables. Ejemplo: detectar el impacto de promociones en ventas, mediante gráficos interactivos. | <https://www.freepik.es/vector-gratis/grafico-vectorial-redes-sociales_756779.htm#fromView=search&page=1&position=1&uuid=c36661b2-0db7-4b0c-a3ac-e9d7f7b7d16a> |
| **Base para otros análisis** | Sirve como entrada para analítica predictiva y prescriptiva. Por ejemplo, utilizar ventas históricas para ajustar inventarios o proyectar la demanda futura. | <https://www.freepik.es/vector-gratis/ilustracion-infografica-analitica_1217965.htm#fromView=search&page=1&position=2&uuid=138a49c0-524a-4713-bc36-45b9815c27c3> |

**Técnicas empleadas**

Las herramientas y métodos aplicados en la analítica descriptiva permiten organizar y presentar datos de forma comprensible, facilitando la toma de decisiones.

1. **Visualización de datos**  
   Representa información numérica y tabulación en gráficos interactivos y *dashboards*. Herramientas como Power BI, Tableau y bibliotecas de Python, como Matplotlib, transforman datos complejos en elementos visuales, fáciles de interpretar.

|  |  |
| --- | --- |
| <https://www.freepik.es/vector-gratis/analista-datos-supervisa-gobierna-ingresos-gastos-lupa-sistema-gestion-financiera-software-finanzas-concepto-herramienta-gestion-ti_11667717.htm#fromView=search&page=1&position=1&uuid=ede6c3e7-2e3a-414e-8fb6-83fa522599ca> | **Ejemplo:** un minorista utiliza Power BI para visualizar las ventas por sucursal en tiempo real, identificando áreas de alto rendimiento y posibles problemas logísticos. |

1. **Estadística descriptiva**  
   Resume datos utilizando medidas como:

Promedio, mediana y moda.

**Tendencia central**

Desviación estándar y rango.  
Estas métricas permiten evaluar ingresos promedios y su variabilidad entre departamentos.

**Dispersión**

Histogramas y diagramas de caja, complementan el análisis.

**Visualización adicional**

1. **Segmentación de datos**  
   Divide la información en categorías o grupos relevantes, mediante algoritmos como K-Means.

|  |  |
| --- | --- |
| <https://www.freepik.es/vector-gratis/dibujado-mano-recopilacion-datos-comerciales_19948901.htm#fromView=search&page=1&position=4&uuid=a07d5828-3175-4a99-aefa-b5cac1e12ada> | **Ejemplo:** una empresa segmenta a sus clientes, según frecuencia y monto de compra, para diseñar campañas de *marketing*, personalizadas. |

**Tabla 1**

*Herramientas y métodos*

|  |  |  |
| --- | --- | --- |
| **Técnica** | **Métodos comunes** | **Ejemplo de uso práctico** |
| **Visualización de datos** | Power BI, Tableau, Matplotlib. | Dashboards para monitorear métricas claves como ingresos y comportamiento del cliente. |
| **Estadística descriptiva** | Medidas de tendencia central y dispersión. | Determinar promedio de ingresos mensuales y variabilidad de gastos en departamentos. |
| **Segmentación de datos** | Agrupación por atributos, algoritmos como K-Means. | Crear estrategias personalizadas para diferentes segmentos de clientes, según su comportamiento. |

La analítica descriptiva organiza y simplifica datos históricos, transformándolos en *insights* prácticos que fundamentan decisiones estratégicas. Su combinación de técnicas como visualización de datos, estadística descriptiva y segmentación, permite a las organizaciones identificar patrones, optimizar recursos y sentar las bases para análisis predictivos y prescriptivos, adaptándose a las necesidades de cada sector.

**1.2. Análisis detallado de la analítica descriptiva**

La analítica descriptiva es una herramienta esencial para las organizaciones, al proporcionar una visión clara y estructurada del comportamiento pasado mediante el análisis de datos históricos. Este enfoque facilita la toma de decisiones informadas, destacando patrones y relaciones que pueden optimizar los procesos operativos y estratégicos. A continuación, se presenta un análisis detallado, combinado con ejemplos, aplicaciones y limitaciones.

**Ejemplo práctico**

Una cadena minorista analiza las tendencias de ventas del último año, utilizando datos desglosados por producto y ubicación, obtenidos de su sistema de punto de venta. Con herramientas como Tableau, se diseñan *dashboards* interactivos que incluyen informes claves sobre productos más vendidos, comparativas regionales y evolución de ventas.

Este enfoque mejora la asignación de recursos y maximiza la capacidad de respuesta al mercado.

Identificación de picos estacionales, como aumentos en ventas navideñas.

Optimización de inventarios para garantizar disponibilidad en los momentos claves.

Creación de estrategias específicas para aumentar las ventas en regiones menos dinámicas.

**Aplicación a la minería de datos**

La minería de datos complementa la analítica descriptiva, al descubrir patrones y relaciones ocultas en grandes volúmenes de información. Un ejemplo destacado es el **algoritmo Apriori**, utilizado para generar reglas de asociación efectivas. El proceso del algoritmo Apriori:

|  |  |
| --- | --- |
| <https://www.freepik.es/vector-gratis/concepto-desarrollo-aplicaciones-telefono-grua_10410996.htm#fromView=search&page=1&position=6&uuid=900f2514-a63e-4d14-ac39-880c0998d59f> | * **Definición de parámetros iniciales:** umbrales de soporte mínimo y confianza mínima, determinan la relevancia de las asociaciones. * **Búsqueda de elementos frecuentes:** se identifican combinaciones de elementos recurrentes en las transacciones. * **Generación de reglas de asociación:** ejemplo: "Si se compra pan, hay un 80 % de probabilidad, que también se compre leche". * **Optimización:** eliminación de combinaciones no frecuentes, para reducir complejidad. |

**Aplicaciones prácticas**

Las aplicaciones prácticas de la analítica, permiten optimizar procesos y tomar decisiones estratégicas en diversos campos.

|  |
| --- |
| **VIDEO** |

**Ventajas de la analítica descriptiva**

La analítica descriptiva ofrece ventajas significativas para las organizaciones, al proporcionar información clara y procesable.

|  |  |  |
| --- | --- | --- |
| **SLIDE** | | |
| **Visualización clara** | Presentación comprensible mediante gráficos y *dashboards*. | <https://www.freepik.es/foto-gratis/reunion-equipo-marketing-reunion-intercambio-informacion-estudiar-actualizaciones_237232435.htm#fromView=search&page=1&position=2&uuid=b186f0c3-5294-45b0-9135-1c54bfc5fa92> |
| **Detección de áreas de mejora** | Identificación de ineficiencias y oportunidades en procesos operativos. | <https://www.freepik.es/foto-gratis/ingenieros-alto-angulo-discutiendo-plan_25125201.htm#fromView=search&page=1&position=15&uuid=9046e448-ad67-4985-b956-945b4deee6b7> |
| **Reducción de incertidumbre** | Apoyo en la planificación estratégica, basada en patrones observados. | <https://www.freepik.es/foto-gratis/profesion-datos-reflejo-lente-simbolo-dinero_1078169.htm#fromView=search&page=1&position=1&uuid=6346d625-f575-4798-ab23-0732dfc03dcc> |
| **Optimización de recursos** | Enfoque en áreas de alto impacto, para maximizar el retorno de la inversión. | <https://www.freepik.es/foto-gratis/concepto-collage-control-calidad-estandar_30589257.htm#fromView=search&page=1&position=7&uuid=aa4602cc-0aab-4c16-850d-5f1342f2cc0a> |

**Limitaciones de la analítica descriptiva**

Aunque la analítica descriptiva es una herramienta valiosa para comprender datos históricos, presenta limitaciones que pueden impactar su utilidad.

|  |  |  |
| --- | --- | --- |
| **ACORDEÓN** | | |
| **Sin capacidad predictiva** | No anticipa eventos futuros ni identifica tendencias emergentes. | <https://www.freepik.es/vector-gratis/eventos-musicales-cancelados_8849702.htm#fromView=search&page=1&position=9&uuid=fc920b16-4e74-4a60-abc0-5f788aa05032> |
| **Falta de recomendaciones específicas** | Requiere que el usuario interprete los resultados, para definir las acciones. | <https://www.freepik.es/vector-gratis/concepto-landing-page-resenas_5156342.htm#fromView=search&page=1&position=0&uuid=bff177ef-6112-46e5-b647-f41ccc6bd8e8> |
| **Dependencia de la calidad de los datos** | Datos incompletos o inconsistentes, pueden llevar a conclusiones erróneas. | <https://www.freepik.es/vector-gratis/personas-que-recopilan-datos-concepto-negocio_20124616.htm#fromView=search&page=1&position=4&uuid=93d7b578-3a88-429a-b047-4183a52dc095> |
| **Limitada profundidad causal** | No profundiza en las causas subyacentes de los fenómenos analizados. | <https://www.freepik.es/vector-gratis/ilustracion-aplanar-tema-curva_8256872.htm#fromView=search&page=1&position=7&uuid=cf2520de-cfc7-4a01-b988-2ed43d40b2d1> |

La analítica descriptiva es el punto de partida para cualquier estrategia de análisis de datos. Su capacidad para organizar y sintetizar grandes volúmenes de información, en visualizaciones claras, la convierte en una herramienta imprescindible para la toma de decisiones. Sin embargo, sus limitaciones destacan la importancia de complementarla con enfoques predictivos y prescriptivos, para profundizar en las causas y proyecciones futuras.

**1.3. Analítica predictiva**

La analítica predictiva aplica modelos estadísticos avanzados y técnicas de *machine learning,* para identificar patrones en los datos históricos, anticipando eventos futuros y proporcionando a las organizaciones, una ventaja competitiva al tomar decisiones estratégicas.

**Técnicas claves**

Las técnicas de analítica predictiva, permiten abordar problemas específicos, como la predicción de valores, clasificación de datos o análisis de relaciones complejas. A continuación, se presentan las más utilizadas, junto con ejemplos prácticos:

**Modelos de regresión**  
Estas herramientas exploran las relaciones entre variables para prever valores o probabilidades.

Utilizada para predecir valores continuos, basándose en variables independientes.

*Ejemplo práctico*: estimar ingresos mensuales de una empresa, en función de su inversión publicitaria.

**Regresión lineal**

Indicada para predecir resultados binarios como "sí/no".

*Ejemplo práctico*: determinar la probabilidad que un cliente cancele un servicio.

**Regresión logística**

**Modelos de clasificación**

Los modelos de clasificación asignan categorías a los datos, según patrones detectados.

**Técnicas de series temporales**

Analizan datos cronológicos para identificar patrones, tendencias y estacionalidades.

***Ejemplo práctico***

Prever la demanda de productos como abrigos en invierno, o helados en verano.

**Técnicas de *machine learning***

Las técnicas de *machine learning* son herramientas poderosas para el análisis y la predicción, en diversos contextos.

|  |  |  |
| --- | --- | --- |
| **SLIDE** | | |
| **K-Nearest Neighbors (KNN)** | Clasifica o predice valores, según datos similares cercanos.   * *Ejemplo práctico*: Identificar clientes con características similares para campañas específicas. | <https://www.freepik.es/vector-gratis/ilustracion-reclutamiento-remoto-dibujado-mano_79310663.htm#fromView=search&page=1&position=4&uuid=4cd5b424-c33d-41f0-bf5d-258964fa4709> |
| **Redes neuronales artificiales** | Modelan relaciones complejas entre datos no lineales.   * *Ejemplo práctico*: Reconocimiento de imágenes en aplicaciones biométricas. | <https://www.freepik.es/vector-gratis/fondo-conexion-red-degradado_12789080.htm#fromView=search&page=1&position=6&uuid=fe10d1d9-2628-4ff9-83cf-933fc2b46cde> |

**Ejemplo práctico: predicción de rotación de clientes**

Una empresa desea anticipar la cancelación de contratos de clientes.

|  |  |  |
| --- | --- | --- |
| **SLIDE** | | |
| **Preparación de datos** | * Recopilación de datos históricos sobre el comportamiento de los clientes.  Limpieza y normalización para garantizar calidad y coherencia. | <https://www.freepik.es/search?format=search&iconType=standard&last_filter=query&last_value=preparacion+de+datos&query=preparacion+de+datos&type=icon> |
| **Modelado** | * Aplicación de regresión logística para predecir probabilidades de abandono.  Uso de Random Forest como validación para mejorar la precisión. | <https://www.freepik.es/icono/profetico_9422881#fromView=search&page=1&position=4&uuid=7c8144ae-e9e8-45ee-b153-320191b094d2> |
| **Resultados y acciones** | * Identificación de clientes en riesgo.  Ofrecimiento de descuentos personalizados y mejoras en el servicio, para aumentar su fidelidad. | <https://www.freepik.es/icono/estadisticas_8713078#fromView=search&page=1&position=9&uuid=234855ea-a0d0-4687-a965-34cef1a820b0> |

**Segmentación previa con K-Means**  
El algoritmo K-Means agrupa a los clientes en categorías, como "clientes fieles" o "clientes en riesgo", según su comportamiento.

**Ventajas de la analítica predictiva**

La analítica predictiva es una herramienta clave para potenciar la estrategia organizacional.

|  |  |  |
| --- | --- | --- |
| **TARJETAS** | | |
| **Anticipación de eventos** | Permite prever comportamientos futuros con base en datos históricos. | <https://www.freepik.es/icono/fiesta_11375624#fromView=search&page=1&position=3&uuid=d20eecc5-8f60-41d1-a91c-f6678ef088b7> |
| **Optimización de recursos** | Focaliza esfuerzos en áreas claves, para maximizar el retorno de inversión. | <https://www.freepik.es/icono/algoritmo_7067974#fromView=search&page=1&position=11&uuid=a5aa9331-e3a2-4435-b545-2ca268f4ed39> |
| **Toma de decisiones estratégicas** | Proporciona información basada en evidencia, para respaldar planes de acción. | <https://www.freepik.es/icono/guia_10547471#fromView=search&page=1&position=8&uuid=e842b82d-d6d6-4901-b934-51e4d5c58c7f> |

Las limitaciones de la analítica predictiva:

Dependencia de la calidad de los datos

Datos incompletos o inconsistentes, reducen la precisión de los modelos.

Requiere conocimientos avanzados

Su implementación y análisis, exigen experiencia técnica y recursos especializados.

La analítica predictiva, transforma datos históricos en información valiosa para el futuro. Combinando técnicas como regresión, clasificación y segmentación, las empresas pueden optimizar decisiones estratégicas, mejorando su capacidad para anticiparse y adaptarse al mercado.

|  |  |
| --- | --- |
| PDF Reader Pro: Editar PDF - Apps en Google Play | **Analítica prescriptiva: guía para optimizar decisiones y procesos.**  Se invita a explorar los fundamentos de la analítica prescriptiva, una herramienta avanzada que transforma datos, en acciones estratégicas. Este enfoque no solo anticipa desafíos futuros, sino que también propone soluciones específicas, para maximizar los resultados. |

**2. Técnicas básicas de minería de datos**

La minería de datos emplea técnicas avanzadas para analizar grandes volúmenes de datos y transformarlos en información útil. Este proceso se fundamenta en la organización y el manejo de datos estructurados y no estructurados, así como en su integración para generar análisis más completos.

**2.1 Datos estructurados**

Los datos estructurados son aquellos organizados en un formato bien definido, como tablas en bases de datos relacionales, donde cada elemento se clasifica en campos predefinidos, con tipos específicos de datos (números, texto, fechas). Estos datos son ideales para su procesamiento y análisis, debido a su estructura rígida y su compatibilidad con herramientas estándar.

|  |  |
| --- | --- |
| <https://www.freepik.es/foto-gratis/empresaria-comiendo-sabroso-sandwich-descanso-comer-trabajando-oficina-empresa-comercial-almuerzo-llevar_15853538.htm#fromView=search&page=1&position=4&uuid=3e165a25-33bc-4498-bd8e-64b205a22014> | **Ejemplo práctico**  En una base de datos de ventas, cada fila representa una transacción, con columnas que indican el ID del producto, la fecha de la venta, el precio, la cantidad y el monto total. Esto permite identificar tendencias de compra, calcular ingresos totales por periodo o analizar el rendimiento de diferentes categorías de productos. |

**Herramientas utilizadas**

El análisis de datos se apoya en diversas herramientas que potencian su eficiencia y alcance.

**Gestores de bases de datos**

MySQL, PostgreSQL y SQL Server, que facilitan el almacenamiento y consulta de grandes volúmenes de datos.

**Lenguajes y bibliotecas**

Python, con herramientas como Pandas y NumPy para manipulación y análisis de datos tabulares.

**2.2. Datos no estructurados**

Los datos no estructurados carecen de un formato predefinido, incluyendo texto libre, imágenes, videos, audios y publicaciones en redes sociales. Este tipo de datos es más complejo de procesar, pero contiene información valiosa que no puede ser capturada por métodos tradicionales.

|  |  |
| --- | --- |
| <https://www.freepik.es/foto-gratis/primer-plano-portatil-iconos-subida_902753.htm#fromView=search&page=1&position=1&uuid=779c0e5a-fa03-4e41-9ed7-848beb6a8326> | **Ejemplo práctico** Un análisis de opiniones de usuarios en redes sociales, como comentarios en Twitter o Facebook, permite identificar percepciones positivas o negativas, sobre un producto o servicio. |

**Herramientas utilizadas**

El manejo y análisis de grandes volúmenes de datos, requieren herramientas especializadas que aseguren eficiencia y escalabilidad

|  |  |  |
| --- | --- | --- |
| SLIDE | | |
| **Hadoop** | A través de su sistema HDFS, almacena datos no estructurados en clústeres distribuidos, asegurando eficiencia y escalabilidad. | <https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcTn77RIGvK0JjIuC1aBdy2wi-SSMmPL0yFhEg&s> |
| **Apache Spark** | Procesa grandes volúmenes de datos en memoria, siendo especialmente útil para análisis en tiempo real. | Apache Spark - Wikipedia, la enciclopedia libre  <https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQk8KLrz1OYfFRDTNayIdRwdanvCy0_Jk8ajg&s> |
| **Procesamiento de lenguaje natural (NLP)** | Técnicas como análisis de sentimientos y extracción de temas recurrentes, en comentarios textuales. | Ilustración de procesamiento del lenguaje natural  <https://www.freepik.es/vector-gratis/ilustracion-procesamiento-lenguaje-natural_21743711.htm#fromView=search&page=1&position=10&uuid=30e742d9-acf7-4157-a608-d49b58af7e11> |

**2.3. Procesos integradores**

La combinación de datos estructurados y no estructurados, es esencial para un análisis integral que abarque, tanto métricas cuantitativas como percepciones cualitativas.

**Uso combinado de datos estructurados y no estructurados**  
La integración permite correlacionar datos transaccionales, con información subjetiva de los clientes.

|  |  |
| --- | --- |
| <https://www.freepik.es/vector-gratis/centro-datos_5039056.htm#fromView=search&page=1&position=2&uuid=239ae08b-2cd8-40bd-a68a-4d3cadeb5ed6> | **Ejemplo práctico**  Si las ventas de un producto muestran una caída, al analizar comentarios de clientes en redes sociales, pueden descubrirse quejas recurrentes sobre su calidad, lo que facilita la identificación del problema y la acción correctiva. |

**Herramientas utilizadas**

La gestión de datos se apoya en herramientas claves que facilitan su integración y análisis.

**ETL (Extract, Transform, Load)**

Estos procesos extraen datos de múltiples fuentes, los transforman en formatos compatibles y los cargan en sistemas centralizados.

**Visualización de datos**

Plataformas como Tableau o Power BI, permiten representar la información combinada en *dashboards* interactivos, que facilitan la interpretación y la toma de decisiones.

**Ejemplo práctico del uso combinado**

|  |  |
| --- | --- |
| Un minorista cruza datos de ventas (estructurados) con comentarios de clientes (no estructurados). Las ventas incluyen columnas como fecha, producto y monto, mientras que los comentarios de clientes, provienen de redes sociales y encuestas. Este análisis revela una correlación entre una disminución en las ventas y menciones negativas en redes, sobre la calidad del producto, lo que permite a la empresa tomar medidas inmediatas. | <https://www.freepik.es/vector-gratis/ilustracion-concepto-extraccion-datos_12325505.htm#fromView=search&page=1&position=7&uuid=239ae08b-2cd8-40bd-a68a-4d3cadeb5ed6> |

**3. Análisis de datos estructurados y no estructurados**

El análisis de datos, tanto estructurados como no estructurados, debe ser evaluado, según criterios claros que aseguren su efectividad, escalabilidad y precisión.

Los criterios para evaluar técnicas de análisis de datos, son:

|  |  |  |
| --- | --- | --- |
| **INFOGRAFÍA INTERACTIVA**  **Análisis de datos estructurados y no estructurados**  Infografía moderna de tres pasos para el flujo de trabajo empresarial  [**https://www.freepik.es/vector-gratis/infografia-moderna-tres-pasos-flujo-trabajo-empresarial\_7550999.htm#fromView=search&page=1&position=23&uuid=67fe16d9-b9ea-4837-80ef-5c04bc7f0b70**](https://www.freepik.es/vector-gratis/infografia-moderna-tres-pasos-flujo-trabajo-empresarial_7550999.htm#fromView=search&page=1&position=23&uuid=67fe16d9-b9ea-4837-80ef-5c04bc7f0b70) | | |
| **Precisión** | Capacidad del modelo para producir resultados correctos y relevantes, medidos por métricas como tasa de error, exactitud y coeficientes de correlación. | <https://www.freepik.es/icono/enfocado_8386247#fromView=search&page=1&position=17&uuid=c8c0e499-0101-40f6-8ea1-82e377ba5c98> |
| **Eficiencia** | Evaluación de tiempo y recursos necesarios para procesar datos, considerando memoria, capacidad de procesamiento y tiempos de ejecución. | <https://www.freepik.es/search?format=search&iconType=standard&last_filter=query&last_value=eficiencia&query=eficiencia&type=icon> |
| **Escalabilidad** | Habilidad del modelo para mantener su rendimiento frente al aumento de datos, utilizando arquitecturas distribuidas o técnicas de optimización. | <https://www.freepik.es/icono/liderazgo_3287070#fromView=search&page=1&position=8&uuid=b366c52b-67a8-4f5d-9777-8dac507a4ca7> |

**Aplicaciones prácticas de los criterios**

La implementación de sistemas tecnológicos avanzados, requiere priorizar atributos claves como la precisión, la eficiencia y la escalabilidad.

|  |  |  |
| --- | --- | --- |
| PESTAÑAS | | |
| Precisión | Asegurar que las sugerencias en un sistema de recomendación de productos, reflejen las preferencias del usuario, basándose en su historial de compras. | <https://www.freepik.es/vector-gratis/fondo-dartboard-objetivo-negocio-objetivo-corporativo_310506666.htm#fromView=search&page=1&position=0&uuid=b7f45e2e-0fc8-45f8-961f-63ad536b2b62> |
| Eficiencia | Procesar datos de transacciones bancarias en tiempo real, para detectar fraudes sin demoras. | <https://www.freepik.es/vector-gratis/fondo-valores-fundamentales-estilo-papel_17748327.htm#fromView=search&page=1&position=9&uuid=a776dbbc-66ba-43d4-bf9a-46687fa36f30> |
| Escalabilidad | Manejar eficientemente millones de datos generados por usuarios nuevos, en una aplicación de *streaming,* sin disminuir la calidad del servicio. | <https://www.freepik.es/vector-gratis/hombre-negocios-que-planea-su-crecimiento-carrera_1311451.htm#fromView=search&page=1&position=1&uuid=652d3771-6d80-4536-87e3-89dab015b1a7> |

Con estas técnicas, herramientas y criterios, la minería de datos permite a las organizaciones, transformar información compleja en decisiones estratégicas, optimizando recursos y resultados.

|  |  |
| --- | --- |
| PDF Reader Pro: Editar PDF - Apps en Google Play | **Ejemplo práctico**  Se invita a profundizar en un caso práctico que compara el análisis descriptivo y predictivo, destacando su implementación con herramientas como SQL y Apache Spark. Este ejemplo ofrece una visión detallada sobre cómo estas metodologías, permiten extraer información clave para la toma de decisiones estratégicas, adaptándose a diferentes necesidades organizacionales y niveles de complejidad. |

**Bibliografía del contenido temático**

Hinojosa Gutiérrez, Á. (2015). *Python paso a paso*. RA-MA Editorial. <https://elibro.net/es/lc/elibrocom/titulos/107213>

Konrad Lorenz University. (s.f.). *Analítica de datos: Tipos y aplicaciones*. <https://www.konradlorenz.edu.co/blog/analitica-de-datos-tipos/>

Navas, J. (s.f.). *Apuntes sobre analítica de datos: Weka y recursos naturales*. Universidad de Jaén. <http://matema.ujaen.es/jnavas/web_recursos/archivos/weka%20master%20recursos%20naturales/apuntesAD.pdf>

Nolasco Valenzuela, J. S. (2018). *Python: aplicaciones prácticas*. RA-MA Editorial. <https://elibro.net/es/lc/elibrocom/titulos/106523>

Peña, S. (2017). *Analítica de datos: Un enfoque integral*. Fundación Universitaria del Área Andina. <https://core.ac.uk/download/pdf/326425169.pdf>

**GLOSARIO DE LA UNIDAD**

|  |  |
| --- | --- |
| **Relacione las definiciones de los términos claves, requeridas para comprender adecuadamente los contenidos de esta unidad. Presentarlo en Orden Alfabético. Máximo 20 palabras.** | |
| **PALABRA, TÉRMINO O ABREVIATURA** | **SIGNIFICADO** |
| **Algoritmo** | Conjunto de reglas y pasos finitos, utilizados para resolver un problema específico o realizar una tarea. |
| **Analítica descriptiva** | Técnica que interpreta datos históricos para identificar patrones y relaciones, proporcionando una base para decisiones estratégicas. |
| **Analítica predictiva** | Uso de modelos estadísticos y *machine learning,* para anticipar comportamientos y eventos futuros, basados en datos históricos. |
| **Analítica prescriptiva** | Método que combina análisis predictivo con simulaciones, para sugerir acciones específicas que optimicen resultados. |
| ***Big Data*** | Conjunto de datos de gran volumen, velocidad y variedad, que requieren tecnologías avanzadas para su procesamiento. |
| ***Dashboards*** | Paneles interactivos que muestran visualizaciones integradas de métricas claves, para facilitar la toma de decisiones. |
| **Datos estructurados** | Información organizada en un formato predefinido, como tablas en bases de datos relacionales. |
| **Datos no estructurados** | Información sin un formato predefinido, como texto, imágenes o videos, que requiere de herramientas especializadas para su análisis. |
| **Minería de datos** | Proceso de explorar grandes volúmenes de datos, para descubrir patrones, relaciones y tendencias útiles. |
| **Tendencias** | Cambios o patrones recurrentes, observados en los datos, a lo largo del tiempo. |

**5. MATERIAL DE APOYO (SIEMPRE SE DEBE INCLUIR AL MENOS DOS RECURSOS BIBILOGRÀFICOS DE APOYO, DOS BASES DE DATOS Y DOS RECURSOS AUDIOVISUALES. SE DEBE INCORPORAR COMO MÍNIMO 2 RECURSOS DE ELIBRO)**

|  |  |  |
| --- | --- | --- |
| **RECURSOS BIBLIOGRÁFICOS DE REVISIÓN BÁSICA**  Incluye libros impresos, digitales, electrónicos, ebook, artículos de revistas impresas, electrónicas, informes, seminarios, congresos, ponencias, presentaciones o diapositivas online, etc. Desde el 2019 en adelante excepto casos excepcionales. Bases de datos (licenciadas por la institución), bases de datos libres, videos bajo licenciamiento creative commons | | **BASES DE DATOS**  Si las referencias bibliográficas son de las bases de datos institucionales mencione únicamente el nombre de la base de datos donde se encuentra el recurso. |
| 1 | Alarcón García, R. E. & Bravo Jaico, J. L. (Dir.). (2021). *Sistema analítico basado en un modelo predictivo de procesamiento de datos en la Big Data en la Educación Superior*. D - Universidad Señor de Sipán. <https://elibro.net/es/lc/elibrocom/titulos/228973> | https://elibro.net/ |
| 2 | Rovira Samblancat, P. & Pascual Cid, V. (2021). *Analítica Visual. Cómo explorar, analizar y comunicar datos*. Anaya Multimedia. <https://elibro.net/es/lc/elibrocom/titulos/217097> | https://elibro.net/ |
| 3 | Charles, V. & Chión, S. (2016). *Analítica de datos para la modelación estructural*. Pearson Educación. <https://elibro.net/es/lc/elibrocom/titulos/107324> | https://elibro.net/ |
| **RECURSOS AUDIOVISUALES**  Incluye videos, podcast, audiolibros, grabaciones sonoras o musicales, que se encuentran alojados en internet. | | |
| 1 | Data\_Pathfinder. (2024). *Creación de un dashboard completo: gestión de datos, medidas y diseño visual* [video]. YouTube.<https://www.youtube.com/playlist?list=PLlOfVkvrHKVQFde9HMXKPB6UKacvVJq2t> | |
| 2 | TED. (2012). *How to get your ideas to spread | Seth Godin* [video]. YouTube.<https://www.youtube.com/watch?v=u_BmGi4RxRQ> | |
| 3 | Aprender Big Data. (2019). (video). YouTube.<https://www.youtube.com/@AprenderBigData> | |
| **PÁGINAS WEB DE REVISIÓN BÁSICA**  Incluye blogs, hipertextos, animaciones, etc. | | |
| 1 | Hernández-Leal, E. J., Duque-Méndez, N. D. & Moreno-Cadavid, J. (2017). Big Data: una exploración de investigaciones, tecnologías y casos de aplicación. *TecnoLógicas*, *20*(39), pp. 15-38. <https://www.redalyc.org/articulo.oa?id=344251476001> | |
| 2 | León García, O. A. (2023). Impacto de las capacidades de análisis de Big Data en la innovación empresarial. *Ingeniería y Competitividad*, *25*(2). <https://www.redalyc.org/articulo.oa?id=291376231005> | |

|  |  |  |
| --- | --- | --- |
| **LABORATORIOS, SIMULADORES, SOFTWARE, APP´S Y HERRAMIENTAS DE IA DE APOYO** | | **LINK DE ACCESO** |
| 1 | Unisys. (s.f.). *Data Analytics Solutions*. | <https://www.googleadservices.com/pagead/aclk?sa=L&ai=DChcSEwi_16CtjaqKAxXKpVoFHWcfEH0YABAGGgJ2dQ&co=1&ase=2&gclid=CjwKCAiAmfq6BhAsEiwAX1jsZyV-GsicmWvyZ4UDt0h6kE8iknIy8JqB1vGMoGKx2ZM3m9wsm1P5pRoCF34QAvD_BwE&ohost=www.google.com&cid=CAESVuD2oLinDIE1KIUkwa71fD4ZE0krJpRhP8OYJsNgLvkJe9cWI5GXYjq8Cht0RTWRGmlwITfTSOIwcaJwD44mc1CHcQpZyPdqQCMrn71AHg4C0KLizV_V&sig=AOD64_1wjiJ-YKlimXyF-eXgT9USMwltkA&q&nis=4&adurl&ved=2ahUKEwiStZutjaqKAxVNTjABHf-CL3YQ0Qx6BAgNEAE> |
| 2 | RapidMiner. (s.f.). *RapidMiner Studio 9.9: Installation guide*. | <https://docs.rapidminer.com/9.9/studio/installation/> |