

# **VERSATILIDAD Y APORTES DEL DATA MINING**

**AUTOR: ALEXANGEL LAZARUS TAPIA**



**San Marcos**

## Versatilidad y aportes del *data mining*

El *data mining* o los *knowledge discovery in databases* (KDD) en el proceso de IN y para su aplicación consiste en acceder, extraer y recopilar de diversas fuentes grupos de data para poder convertirla en información legible, aplicable y manejable; para que pueda ser usada posteriormente. Es importante aclarar, que el *data mining* realmente configura un conjunto de técnicas de extracción de datos para detectar patrones de comportamiento y su soporte son los algoritmos matemáticos. El *data mining* es una fase que forma parte del KDD y requiere el poder realizar análisis de grandes cantidades de datos para establecer relaciones y conectores entre ellos; entonces se puede colegir que la minería de datos incluye tres fases: la selección de objetivos, la selección del algoritmo (o algoritmos) y su uso.

En ese sentido y como herramienta permite ser aplicada con diferentes perspectivas; sin embargo, se listan las más usuales:

- **Predicción automática de comportamientos:** se puede lograr resolviendo problemas inicialmente de clasificación; en el caso financiero, un banco a partir de un estudio basado en la minería de datos puede evaluar los resultados de las campañas de mercadeo realizadas anteriormente con el ánimo de identificar patrones de comportamiento entre sus usuarios y relacionarlos con el consumo de sus productos, preferencias, hábitos de compra y pago; de este modo se pueden configurar grupos de clientes con los mismos perfiles y hábitos de consumo, para programas de mercadeo puntuales, para racionalizar los recursos del área y de la empresa.
- **Predicción automática de tendencias:** se soporta en bases de datos históricas, el sistema permite crear un modelo para predecir las tendencias. Por ejemplo, podemos pronosticar las ventas en el futuro o realizar también pronósticos sobre la demanda de música de un género en particular o pronósticos sofisticados sobre el comportamiento de divisas o del mercado bursátil.
- **Descubrimiento automático de comportamientos desconocidos anteriormente:** el sistema permite la visualización y *clustering*, además de tratar los datos empresariales bajo una perspectiva diferente, para así establecer relaciones entre estos.



### Clustering

Técnica de minería de datos (data mining) dentro de la disciplina de Inteligencia Artificial que identifica de forma automática agrupaciones o clústeres de elementos de acuerdo a una medida de similitud entre ellos. El objetivo fundamental de las técnicas de clustering consiste en identificar grupos o clústeres de elementos tal que:

La similitud media entre elementos del mismo clúster sea alta. Similitud intraclúster alta.

La similitud media entre elementos de distintos clústeres sea baja. Similitud inter-clúster baja.

La identificación de clústeres o grupos de elementos se basa en una medida de similitud. Diferentes medidas de similitud dan lugar a diferentes clústeres.

---

## Usos del *data mining*

*Clustering analysis* **para identificar grupos objetivo:** este recurso aglutina a partir de archivos, las cualidades y características afines a un grupo poblacional en particular. Estas características pueden ser el estrato, ubicación, actividad, formación, ingresos, consumo, etcétera. Se trata de una técnica de minería de datos que le facilita al departamento de mercadeo realizar microsegmentaciones desde las bases de datos para el diseño de nuevos productos, para el diseño de promociones, programas de fideliza-

ción, estrategias de mercadeo directo, referidos y planes de fidelización. Todas estas posibilidades mediadas por la agrupación determinada por las similitudes y combinaciones que se puedan presentar. Las combinaciones no tienen límite, más bien deben responder en su selección a las implicaciones de cualidades y características necesarias y exigibles para lograr la configuración de un clúster que sea efectivo.

**Regression analysis para hacer proyecciones de mercadeo:** prever el futuro es la pretensión del mercadeo, pero el aporte del análisis de regresión, como técnica de *data mining* permite evaluar la evolución en el consumo, variación en los hábitos, comportamiento de las tendencias en la favorabilidad de marcas en una categoría en particular y nuevas expectativas del consumidor que inciden en los presupuestos para campañas publicitarias y planes de acción relacionados con el mercadeo y la gestión comercial.

**Classification analysis para identificar spam y más:** da soporte a las interacciones entre empresa y clientes de acuerdo con el despliegue del plan de mercadeo digital y de CRM, clasificando la comunicación *in bond/ out bond*, los efectos de las campañas y la relación entre el mercado actual y potencial a partir del impacto

que estas producen. En consecuencia, este procedimiento permite identificar esquemas repetitivos o patrones que se presentan en el *data mining* y su clasificación, en sintonía con las variables que se incluyan en esta técnica.

**Anomaly detection para reconocer las incongruencias:** un problema medular de la administración de datos y la información en cuanto a su transparencia, calidad y fiabilidad tiene que ver con el

usuario o posible manipulación anómala que provoca desvíos, lecturas que no corresponden y por ende errores que pueden resultar onerosos para la empresa. Para su corrección o eliminación frente a la cualidad y estandarización de lo recopilado, frente a desvíos y contenidos incongruentes o posibles inconsistencias en la base de datos se utiliza la herramienta de *anomaly detection*, que está facultada para detectar y hallar lo referido que se presente en el *software*, “limpiando” lo incorrecto pues el recurso está habilitado para realizar procesos sofisticados en bases de datos que pueden contener N registros, que manualmente serían imposibles de afinar.

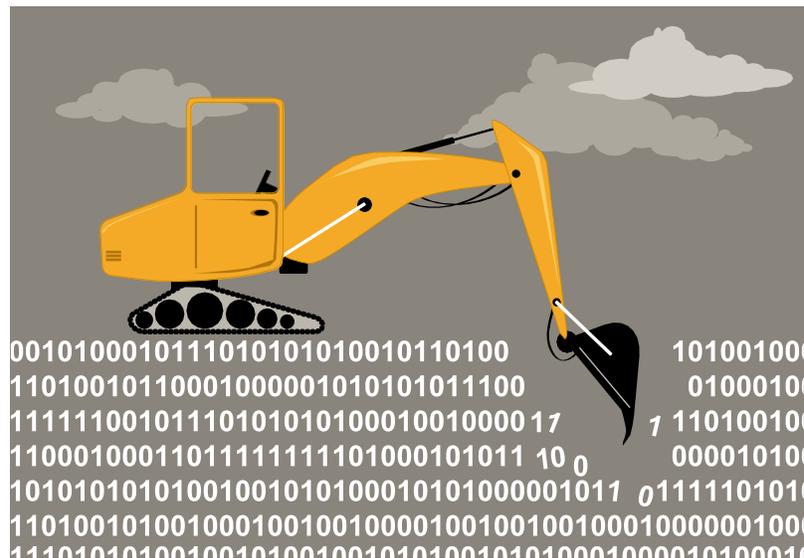


Figura 4.

Fuente: Shutterstock/513585007

**Intrusion detection para una mayor seguridad del sistema:** los procesos de mercadeo y la información que recauda esa área y la empresa en general deben ser sujetos de protocolos de seguridad que van más allá de lo legal; pero por las exigencia y posibilidades del manejo de información multidimensional, tanto mercadeo como seguridad, deberían ser dos conceptos

manejados al unísono, son complementarios; el beneficio que otorga esa complementariedad radica en que las acciones y estrategias de mercadeo que se soportan en datos o información contaminada, por su mal manejo, se convierte en inservible o es inductora de errores. En ese aspecto, esta herramienta identifica archivos y fuentes desconocidas, purifica las bases de datos y provee de mayor seguridad a todo el sistema.

**Association rule learning para descubrir la relación entre datos:** su uso está dado y se hace más útil cuando se trata de empresas que producen y comercializan diferentes líneas de productos y en grandes cantidades, sin importar el modelo de comercialización y distribución (virtual o física) y su cobertura geográfica, indistintamente del canal y la modalidad, pues su potencial de uso permite implementar relaciones e interacciones con la multitud de data recogida y almacenada; por ejemplo, si una plataforma digital ofrece como eje de negocio la transacción virtual con apoyo logístico real, este recurso permite establecer las cantidades de compradores que compran en línea un producto; si el diseño y navegación por la web estimula la compra de otro así no sea complementario, en ese caso tan simple se pueden tomar decisiones sobre paquetes promocionales, ajustar inventarios, disminuir costos logísticos, realizar pronósticos de producción y compras, etcétera.

**Decision trees para optimizar el project risk management:** la complejidad en la toma de decisiones contiene naturalmente riesgos y es en esa medida que el sistema puede presentar la alternativa de un árbol de decisión, la herramienta jerarquiza y establece relaciones que facilitan las decisiones, evidencia los probables ries-

gos en un escenario previsto, aporta a la disminución de costos, da una visión a la gerencia sobre las irregularidades o variaciones que impliquen incertidumbre; por lo tanto, el mismo árbol se convierte en una valiosa herramienta para el *project risk management* (gestión del riesgo).

**Neural networks para automatizar el aprendizaje:** cuando se cohesiona el *clustering* y se fusiona con los árboles de decisión, se puede hablar de la configuración de una red neuronal, destacando que es parte de los avances de la minería de datos y esta herramienta formula relaciones y como consecuencia de su aplicación se da un nuevo conocimiento mucho más complejo y sofisticado, dado que genera mayor profundidad y en consecuencia más variados usos, dado que la red neuronal amplía ese conocimiento y lo hace más preciso, lo que lleva a dar soporte a las estrategias y acciones que se pretenda implementar.

**Rule induction para análisis predictivos basados en los datos:** este recurso identifica comportamientos recurrentes y los estandariza, lo mismo que los irregulares; los resultados apoyan el planteamiento predictivo a partir de los mismos pronósticos, cuyo elemento natural es el número infinito de registros que conforman las bases de datos; además al detectar estos comportamientos ayuda a reaccionar y más aún anticiparse a la competencia con mínimos márgenes de error.

De otro lado, dentro de los contenidos que forman parte del proceso de IN, la multidimensionalidad permite graficar y hacer comprensibles los datos para que el usuario pueda acceder e interpretar los mismos y así estar en capacidad de atender problemas contenidos en los sistemas organizacionales.

Las fuentes son amplias y son desde un Excel, bases de datos convencionales, gráficas, estadísticas, etcétera, destacando que la aplicación de estas herramientas se hace tangible desde la ubicación por la dispersión de la información, facilitar la accesibilidad, flexibilidad para adaptarla según requerimientos y como elemento unificador, el análisis que se traducirá en conocimiento.

La multidimensionalidad entonces permite incluir datos sobre atributos y variables de procedencia diversa, tal y como se expone a continuación.

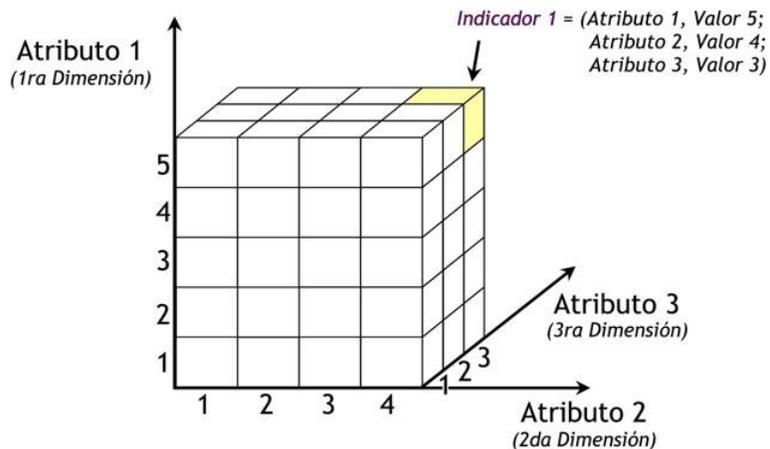


Figura 5. Cubo multidimensional M  
Fuente: dataprix.com

## La aplicación analítica

La aplicación analítica no es solo un menú de informes creados aleatoriamente, sino que tiene unas cualidades derivadas de la intención del estratega, con la pertinencia de ser una aplicación de negocios que posee una lógica de uso, mediada por un menú de información para su modificación y ajustes, que propende por el acceso y manejo fácil por parte de usuarios disímiles —sin importar su perfil— para evaluar y tomar decisiones, que impacten en el incremento de eficiencia y aprendizaje para un departamento en particular, pero de manera más ambiciosa para las unidades de negocio de la empresa y su nivel de competitividad.

Las características exigibles para que funcione la aplicación analítica son las siguientes:

**Integradas lógicamente:** las herramientas concurrentes en la IN son aplicaciones que son susceptibles de permitir el acceso y manejo de modo amigable por parte del usuario, además de sus posibilidades de ejecución y servicio en todos los niveles de la organización, dado lo sistémico que se exige en el proceso; es decir, las relaciones causa-efecto, que se traducen en conocimiento fino y complejo. Por ejemplo, empleados que por su rol dentro de la organización requieren de información actualizada, amerita el soporte de guías analíticas para que el manejo de los datos sea veloz, fiable y como soporte en la gestión estratégica y operativa.

**Informes interactivos:** exigen versatilidad, amigabilidad, accesibilidad y confiabilidad para realizar la gestión del usuario que los necesita; por lo tanto, no deben ser planos ni el usuario un mero lector, sino más bien su funcionalidad está dada en la facultad que tiene el usuario de modificar, incluir y ajustar contenidos sin desconocer los protocolos y naturaleza de la fuente o del sistema mismo; esta versatilidad apoya el análisis desde la complejidad de la información. Además, los resultados pueden ser surtidos a otros niveles más complejos con la flexibilidad de administrar la información y lo que esto implica. Estos informes que tienen la condición de ser interactivos se pueden emanar con múltiples formatos, según la plataforma y *software* de la organización, como por ejemplo en [cubos OLAP](#), técnicas avanzadas de visualización y formulación estadística.



#### Cubos OLAP

Estructuras multidimensionales (cubos) que permiten analizar bases de datos relacionales de gran volumen y variedad con una gran agilidad y rapidez, reduciendo enormemente el tiempo y los recursos empleados en el análisis.

---

**Información integrada:** generalmente aglutina contenidos de varias fuentes y su uso es el análisis, se constituye en un repositorio de almacenamiento de datos de la empresa bajo el entendido que un almacén de datos tiene el potencial de contener amplias aplicaciones analíticas.

**Dirección de un dominio comercial:** las aplicaciones analíticas tienen que ver con las necesidades de información para su diseño, además de su confrontación con el sector en el que se va a implantar la solución innovadora; permiten evidenciar el comportamiento de la gestión de ventas, las inversiones y las estructuras de capital de empresas en curso, aspectos relevantes del mercado. Así mismo, la aplicación puede responder frente a necesidades de detección de oportunidades comerciales, para nuevos productos, aparición y tendencias que configuren nuevos nichos de mercado, la variación y evolución de la demanda, las ventas de marcas por líneas y familias, el comportamiento de la demanda por sectores y regiones; etcétera.

Los elementos que configuran una aplicación analítica son básicamente seis:

- **Visualización y percepción:** es la presentación amigable y gráfica que permite lecturas amplias, que junto a la percepción que tiene que ver con la navegación por los sistemas de manera comprensible y accesible, en los diferentes departamentos de la empresa o por tipos de usuario.
- **Reportes y análisis:** de acuerdo con el diseño y configuración de los sistemas, los documentos y reportes tienen dos opciones genéricas: aquellos que solo registran información pero no permiten su manipulación en términos de contenidos y los que facultan al usuario a modificar e incluir contenidos; asumiendo que los contenidos tienen que ver en su despliegue con cálculos prolijos, algoritmos estadísticos y visualizaciones de última generación para la administración de cantidades ilimitadas de información.
- **El servidor de análisis:** este recurso permite procesar requerimientos del usuario teniendo en cuenta el run-time de metadatos previstos y archivarlos en el almacén del servidor.



#### Run-time

Tiempo de ejecución al intervalo de tiempo en el que un programa de computadora se ejecuta en un sistema operativo.

---

- **Lógica de negocios:** es un tipo de lógica aplicable para la empresa, incluye las reglas de navegación de usuarios, reportes, accesos, protocolos, restricciones y ejecuciones.
- **Modelo de datos:** es la médula de las herramientas que surten análisis, imprimiéndole transversalidad a su uso en la IN; los parámetros y esquemas de manejo de datos son los que incorporan las soluciones de los cubos de OLAP.
- **Mapeo de las fuentes:** se definen como la trazabilidad y la decisión de realizar ajustes para gestionar los datos fuente, dentro del modelo de datos designado.

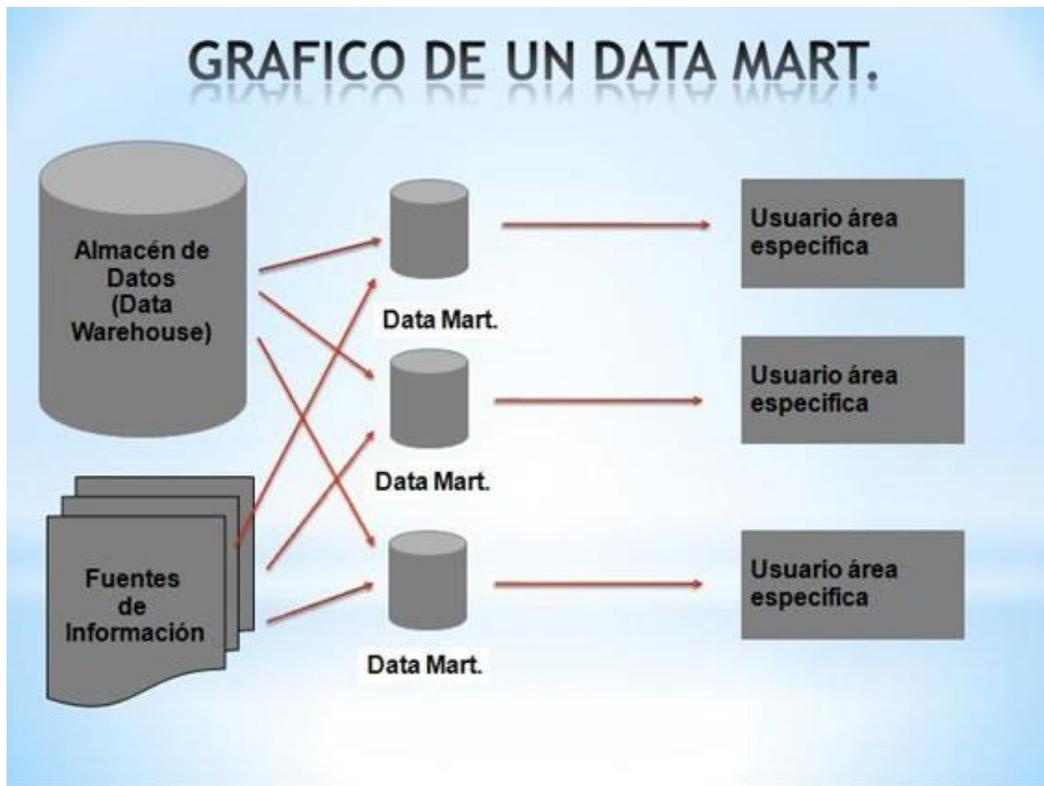


Figura 6. El *datamart*: funcionalidad

Fuente: <https://sites.google.com/site/itsginteligenciadenegocios/home/1-1-conceptos-basicos/1-1-2-data-mart>

Se define como la especialidad en la administración y uso de una base de datos que particularmente es requerida de manera habitual por área o departamento de la empresa, lo que permite inferir que su contenido es específico, puesto que el almacenamiento de los datos sirve de manera enfocada para ese tipo de usuario que le otorga su utilidad e interpretación, por lo tanto sus contenidos son precisos, vigentes y deben garantizar el espectro necesario para el fin que están concebidos; además, se relaciona su particularidad mediante la misma funcionalidad del área que los usa, incidiendo en la gestión y resultados de la misma. Un *datamart* puede ser alimentado desde los datos de un *data warehouse* o integrar por sí mismo un compendio de distintas fuentes de información.

En su diseño; el *datamart* de un área funcional de la empresa parte de precisar cuál debe ser su estructura ideal para facilitar el análisis de la información, enfatizando que esta puede incorporarse sobre una base de datos OLTP, como el propio *data warehouse*, o sobre una base de datos OLAP; el modelo se determinará por el tipo y cantidad de datos, los requerimientos y las particularidades que declare el usuario en sincronía con el modelo de negocio, estructura organizacional, modelo de procesos, sistemas de información de la empresa, etcétera. En ese sentido, se tipifican dos alternativas de *datamarts*:

- **Datamart OLAP:** se basan en los populares cubos OLAP, que se construyen agregando, según los requerimientos de las áreas, las dimensiones y los indicadores necesarios de cada cubo relacional. El modo de creación, explotación y mantenimiento de los cubos OLAP es muy heterogéneo, en función de la herramienta final que se utilice.
- **Datamart OLTP:** se basan en un simple extracto del *data warehouse*, pero esto no inhibe la introducción de mejoras a partir de las particularidades de cada área. Las configuraciones que más se utilizan son las tablas *report*, que vienen a ser *fact-tables* reducidas (que agregan las dimensiones oportunas), y las vistas materializadas, que se diseñan e implementan con similares insumos y procedimientos que las anteriores, pero con un propósito central: explotar la reescritura de *queries*.



### Fact-table

Tabla de hecho es la tabla central de un modelo en estrella. Por ejemplo, son “hechos” las ventas, los pedidos, los envíos, las reclamaciones, las compras, etc. Es decir, son todas aquellas medidas numéricas que incluiremos en nuestro sistema IN.

### Query

Hace referencia al término o concepto que escribimos en Google u otros buscadores cuando realizamos una búsqueda por palabra clave o keyword.

Dicha consulta dará como resultado una SERP, que es el acrónimo de Search Engine Result Page (en español, página de resultados de búsqueda). Esto es, ni más ni menos, que la página resultante de una búsqueda en Google con sus diversas entradas.

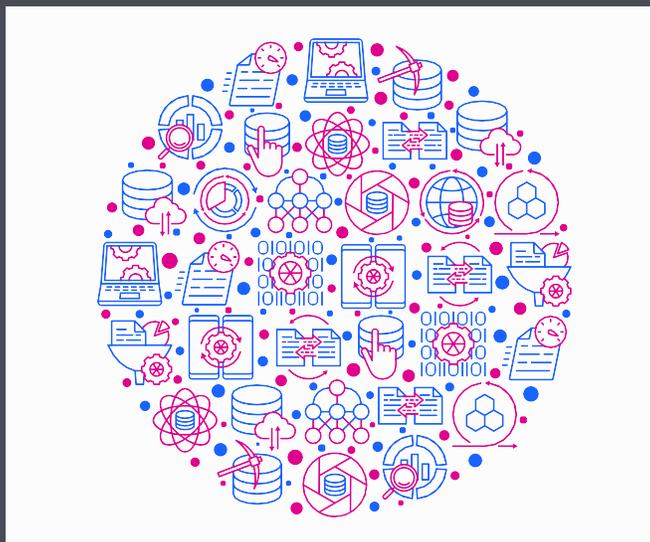


Figura 7.  
Fuente: Shutterstock/1079404394



## Instrucción

A continuación, lo invitamos a ampliar sus conocimientos respecto a este tema con la actividad de aprendizaje: control de lectura, que encontrará disponible en la plataforma.

## Los DSS (*decision support system*)

Conforman un sistema de información que es flexible y permite la interactividad, ayudando a quien toma las decisiones para que utilice contenidos y técnicas para la solución de problemas estructurados o semiestructurados. Esta herramienta le sirve a la alta gerencia o al estratega en el diseño e implementación de estrategias y la toma de decisiones, puesto que suministra a partir de un análisis sistemático y sistémico, varias alternativas o escenarios que incluyen variables de alto impacto y de riesgo, mediado lo anterior por modelos y herramientas tecnológicas. Sin embargo, es pertinente aclarar que esta herramienta no soluciona problemas, más bien se convierte en una forma de evaluar opciones y realizar análisis antes de tomar cualquier decisión trascendental para la organización.

Gerencialmente se puede afirmar entonces que el principal objetivo de los **sistemas de soporte a decisiones** es usar de manera estratégica todo el contenido que se concentra en una base de datos corporativa (*data warehouse* o *datamart*) pues en comparación con otras soluciones tales como el cuadro de mando (**CMI**) o los sistemas de información ejecutiva (EIS), que son posteriores al planteamiento estratégico; la herramienta detallada permite el uso de información versátil y flexible, además de su interacción y fácil navegación, incluyendo elementos amigables, estéticamente armoniosos y legibles, a la vez que sencillos para su comprensión por parte de diferentes usuarios. Otra cualidad de esta herramienta es que se puede extrapolar a cualquier nivel jerárquico dentro de la empresa, en problemas de diverso orden (estructurados como no estructurados), mientras que la solución CMI es excluyente y generalmente su uso se da en la alta dirección.



### Cuadro de mando integral (CMI)

Modelo de gestión que traduce la estrategia en objetivos relacionados entre sí, medidos a través de indicadores y ligados a unos planes de acción que permiten alinear el comportamiento de los miembros de la organización con la estrategia de la empresa.

Se trata de una herramienta de control de gestión, cuya función primordial es la implantación y comunicación de la estrategia a toda la empresa.

Las características y beneficios más significativos del DSS son

- Informes dinámicos, flexibles e interactivos que permiten al usuario opciones diferentes a los menús preestablecidos que se tuvieron en cuenta al tiempo de la implantación, pero sin tener en cuenta las necesidades de usuarios.
- No requiere conocimientos técnicos, permite diseñar nuevos objetos icónicos, gráficas, informes y reportes, es posible su navegación sin restricciones ni habilidades técnicas o de expertos.
- Rapidez en la generación de resultados o respuesta, su soporte se encuentra en la base de datos que es parte de la constitución del *data warehouse* corporativo o en área (*datamart*). Además, en su diseño incorpora modelos de datos en estrella o copo de nieve, aunado a que en su estructura es posible la administración y almacenamiento de altos volúmenes de información.
- Permite la sinergia y la integración entre todos los sistemas/departamentos de la compañía.
- Cada usuario puede acceder a la información pertinente y de acuerdo con sus requerimientos; además, una bondad de la herramienta es delimitar sin restringir el acceso a la información relevante, elevando la eficiencia y evitando la saturación de información no requerida para el sistema ni el usuario.
- Disponibilidad de información histórica. El DSS permite confrontar los datos actuales con información histórica con el fin de determinar ciclos, patrones y comportamiento; analizar tendencias, y evaluar la evolución del comportamiento del negocio, entre otros asuntos.



#### Instrucción

En este punto, lo invitamos a desarrollar la actividad de aprendizaje: demostración de roles. Disponible en la plataforma.

## Los sistemas de información

### Tipos de Sistemas de Información



Figura 3.  
Fuente: <https://bit.ly/2QSpwX1>

Los sistemas de información adquieren relevancia superlativa en la medida que contribuyen al incremento de la capacidad organizacional ante la incertidumbre y frente a los cambios intempestivos del contexto del negocio y en el caso del emprendedor para validar la factibilidad de su proyecto de negocios; en ese escenario cambiante, se convierte en una necesidad apremiante para tomar decisiones rápidamente y además contrarrestar la agresividad del entorno, para ello se torna necesaria la inmediatez de la información procesada, confiable y vigente.

Los sistemas de información son considerados como una herramienta gerencial que fomenta el acceso oportuno de información relevante, que acompañe la toma de decisiones, insistiendo en la gestión y análisis de información relevante, de manera oportuna y ágil, en un ambiente de incertidumbre, para mitigar y disminuir riesgos y errores.

La toma de decisiones se apoya en los SI de la siguiente manera: 1) sirve para la gerencia diaria, para apoyo a la operación del negocio y gestión empresarial; 2) aporta información valiosa y de nivel para las estrategias y toma de decisiones, y 3) la determinación y como inductor de para la construcción de la ventaja competitiva, estimulando su implementación y explotación.

Para lograr estos planteamientos es menester considerar estos dos tipos de sistemas: expertos (aportan en el apoyo a actividades y operaciones empresariales) y los de gestión estratégica (proveen a la organización de ventajas estratégicas frente a sobre las fuerzas competitivas inherentes a la globalización).

El nivel de importancia que tiene el uso de los SI en la toma de decisiones se hace tangible en el acceso a la información necesaria y fiable para que la gerencia o el estratega marque el rumbo competitivo de la empresa mediante las decisiones que mejoren la posición competitiva, pero al interior de la organización produce beneficios establecidos en incrementos de la eficiencia operativa y en la gestión. De manera transversal, impacta en las competencias del talento humano, permite el análisis actualizado sobre factores críticos relacionados con la generación de valor y adicionalmente propone improntas para enfrentar los desafíos ya no nacionales sino globales.



#### Instrucción

Para finalizar el desarrollo temático de este eje, lo invitamos a recopilar los principales aspectos abordados por medio del recurso de aprendizaje: video-resumen, que encontrará en la plataforma.

## Reflexiones

Para un proyecto innovador es importante incorporar la IN pues en su aplicación se encontrarán las mejores alternativas para que la toma de decisiones, la fijación de objetivos y el cálculo de riesgos, permita su implementación con improntas futuribles de competitividad y capacidades de generación de valor. Es pertinente entonces tener en cuenta las siguientes consideraciones para su aplicación:

- Los datos y la información permiten mejoras continuas y hacen visibles los resultados obtenidos en las estrategias aplicadas, basándose en términos de precisión y alcance.
- Las variables, fuentes y recursos a implementar en la IN son parte de la decisión estratégica sobre el qué analizar, dónde está la información relevante, qué herramientas utilizar, quiénes son los que administran la información, quiénes los usuarios y cuáles los propósitos del proceso.
- Representar formalmente medidas, modelos y los patrones más significativos de datos a gran escala puede lograrse con diversas fuentes, herramientas y métodos.
- Aplicar la IN es una exigencia empresarial y para los innovadores, dado que al ser un proceso sistemático y sistémico de búsqueda, selección, análisis de información relevante y pertinente, que ausculta condiciones del entorno y del sector en el que opera la empresa o se va a implantar la innovación, conduce a las posibilidades de análisis que apoyen la toma de decisiones y cuyo fin último se tipifica en los siguientes usos:
  - Genera alertas tempranas que permiten anticiparse a condiciones favorables o desfavorables del entorno.
  - Cataliza a la acción: promueve cambios en la organización y en la visión personal del emprendedor.
  - Crea beneficios dado que fomenta las estrategias hacia la consolidación de la posición competitiva de la empresa o el impacto que puede producir la solución innovadora para un emprendedor.
- Es prudente destacar y marcar las diferencias entre la IN y el *big data*, dado que pueden conducir a confusiones porque se trata del uso de tecnologías con un propósito aparentemente idéntico: el análisis de datos. Sin embargo, existen algunas diferencias entre estas dos tecnologías, tanto en su operatividad como en el tipo de datos que analizan. *Big data* va un paso más allá de la IN, pero sin ella no se requiere *big data*; por lo tanto, se permite aseverar que la IN es el paso preliminar al *big data*.

- Las principales diferencias entre negocio e inteligencia, *big data* y sus alcances son que la IN tiene como objetivo el análisis de datos consolidados en un entorno llamado *data warehouse*, estructurados en una base de datos relacional convencional con un conjunto adicional de índices y formas de acceso a las tablas, vistas, cubos multidimensionales o modelo tubular; mientras que en el *big data* los datos se almacenan en un sistema de ficheros distribuido, en lugar de en un servidor central.
- Descubrir los patrones de comportamiento de los usuarios potenciales se convierte en soporte para las estrategias de mercadeo y ventas.
- La administración de la información convertida en conocimiento es hoy por hoy un componente esencial de la ventaja competitiva de cualquier empresa y genera valor para los accionistas.
- Es imperativo para la empresa y para el emprendedor descubrir nuevas alternativas tecnológicas accesibles que les permitan el acceso a datos e información, para tomar decisiones de alto contenido, disminuyendo el riesgo a partir de la calidad y confiabilidad de la información analizada.

# BIBLIOGRAFÍA

Ahumada, E., & Perusquia, J. M. A. (2016). Inteligencia de negocios: estrategias para el desarrollo de competitividad en empresas de base tecnológica. *Contaduría y Administración*, 61, 127-158.

Bigdatasipb. (2017). Cubos OLAP. Recuperado de <https://bigdatasipb.wordpress.com/2017/08/30/cubos-olap/>

CMlgestión (s. f.). Cuadro de mando integral. Recuperado de <http://cmigestion.es/cuadro-de-mando-integral/>

Conexión ESAN (2015). Las redes neuronales en la inteligencia de negocios. Recuperado de <https://www.esan.edu.pe/conexion/actualidad/2015/06/11/redes-neuronales-inteligencia-negocios/>

Howson, C. (2009). *Business intelligence: estrategias para una implementación exitosa*. México D.F., México: Mc-Graw-Hill Interamericana.

Romero-Campero, F. J. (s. f.). Búsqueda de patrones: técnicas de *clustering*. Recuperado de [https://www.cs.us.es/~fran/curso\\_unia/clustering.html](https://www.cs.us.es/~fran/curso_unia/clustering.html)

Rouse, M. (2017). Inteligencia artificial o AI. Recuperado de <https://searchdatacenter.techtarget.com/es/definicion/Inteligencia-artificial-o-AI>





[www.usanmarcos.ac.cr](http://www.usanmarcos.ac.cr)

San José, Costa Rica