**Módulo 4: Tecnología, ciencia e inteligencia**

**Introducción:**

¿Te has parado a pensar en la cantidad de cosas que hacen las máquinas por nosotros? Por máquinas no solo nos referimos a máquinas grandes, sino incluso a máquinas muy pequeñas e incluso invisibles a nuestros ojos. Algoritmos que filtran, por ejemplo, los mails que recibimos en nuestra bandeja de entrada y que limitan los ciberataques. Pero ¿qué puedes decir de los robots aspiradores que te limpian el suelo cuando detectan suciedad?, ¿o los avances en salud, como los sensores que implantan en el brazo a personas diabéticas para monitorear el azúcar en sangre y avisarles de la necesidad de insulina a través del móvil? ¿Sabes qué es todo esto? Inteligencia Artificial.

La Inteligencia Artificial ya ha dado el salto de la pantalla a la realidad. Un salto que se convertirá en una revolución equiparable a la que generó Internet. ¿Qué es realmente la Inteligencia Artificial? Una combinación de algoritmos y máquinas que replican las capacidades humanas. Una tecnología que por un lado parece lejana y misteriosa, y por otro ya forma parte de nuestra rutina.

Pero la inteligencia artificial nos afecta indirectamente en muchos más aspectos. Al impulsar el big data y procesar grandes cantidades de datos, ofrece numerosas ventajas comunicacionales, comerciales y empresariales. En finanzas, las tecnologías inteligentes pueden ayudar a los bancos a detectar el fraude, predecir patrones del mercado y aconsejar operaciones a sus clientes.

Para el ámbito comercial, empresas como Amazon utilizan robots para identificar si un libro tendrá o no éxito, incluso antes de su lanzamiento. Entre los avances agrícolas existen plataformas específicas que, por medio de análisis predictivos, mejoran los rendimientos agrícolas y advierten de impactos ambientales adversos.

No hay nueva tecnología sin incertidumbre. Los beneficios de la Inteligencia Artificial se enfrentan a sus potenciales peligros. Todo está en manos de la huella que dejemos los humanos con su uso. Los avances en robótica e inteligencia artificial supondrán un cambio en todos los aspectos de nuestra vida. ¿Positivos o negativos? Solo el tiempo lo dirá.

**Objetivo:**

Comprender las bases de la inteligencia artificial y el manejo de datos, con el fin de entender la ciencia de datos y conocer el uso de indicadores en los negocios de una empresa.

**4.1. Inteligencia artificial y manejo de datos**

#### Introducción:

La tecnología se ha movido de la ciencia ficción al mundo empresarial, para enfrentar los desafíos actuales y prepararse para el futuro, por esto se requieren soluciones de Inteligencia Artificial que se sumen a las estructuras y estrategia de los datos. La infraestructura de I.A. permite pasar de un programa piloto a una estrategia completa de integridad de los datos. Las organizaciones que consiguen aplicar con éxito la inteligencia artificial a escala adquieren un rendimiento financiero superior.

Permite dar solución a un gran volumen de datos disponibles para guardar, analizar, procesar y crear valor y conocimiento. Las bases de datos estructuradas han dado respuesta a esta situación, ya que el reto actual y del futuro es combinar este tipo de datos estructurados con otros datos llamados no estructurados, estos son datos extraídos de distintos dispositivos, aplicaciones o formatos. Sin duda, un universo complejo de datos que, uno a uno, tienen un valor importante, pero que combinados multiplica ese valor.

Tener una [visión clara de la necesidad del cliente](https://www.apd.es/atencion-al-cliente-digital/) es más fácil que nunca. El Big Data permite recopilar información de las visitas de la página web, redes sociales, registros de llamadas y otros orígenes de datos, para perfeccionar la experiencia de interacción y optimizar el valor que se ofrece; disminuir las tasas de abandono de los visitantes, gestionar incidencias de forma proactiva, todo esto con un único propósito: ganar más clientes.

La recolección de datos puede ayudar a innovar en las organizaciones gracias al análisis de las interdependencias entre instituciones, procesos y entidades. Las posibilidades son ilimitadas: mejorar la toma de decisiones financieras y las consideraciones de planificación, implantación de precios dinámicos, análisis de tendencias y los deseos de los clientes para desarrollar nuevos servicios y productos.

#### Objetivo:

Demostrar que la IA simplifica la infraestructura de almacenamiento, en etapas de transición, con el fin de obtener mejor información, con mayor valor y mayor ventaja competitiva.

Los temas a revisar para responder al objetivo de esta sesión son:

* Inteligencia artificial, Big data, Minería de datos y Lenguaje de máquina
* Funciones y alcance de los datos desde un enfoque científico

#### Big Data

El sector empresarial presenta gran desconocimiento sobre lo que significa Big Data; hoy las compañías no saben qué hacer con el gran volumen de datos e información almacenada en diferentes medios o bases de datos, los cuales pueden ser de gran importancia, principalmente en la toma de decisiones.

Como un problema de Big Data se puede contemplar la forma como hoy crecen los datos en volumen, velocidad y variedad; esto es debido al gran avance y uso de las tecnologías de información, y al uso diario que las personas hacen de ellas.

#### Minería de datos

El rápido crecimiento del interés en la minería de datos es debido:

1. Al avance de la tecnología de Internet y a la gran participación en aplicaciones multimedia en este dominio,
2. A la facilidad en la captura de datos y el abaratamiento de su almacenaje,
3. A compartir y distribuir los datos en la red, junto con el aumento de nuevas bases de datos en los repositorios,
4. Al desarrollo de algoritmos de aprendizaje automático robustos y eficientes para procesar estos datos,
5. Al avance de las arquitecturas de las computadoras y la caída del coste del poder computacional, permitiendo utilizar métodos computacionalmente intensivos para el análisis de datos,
6. La falta de adaptación de los métodos de análisis y consulta convencionales a nuevas formas de interacción y finalmente,
7. A la potencia que este tipo de análisis vienen mostrando como herramientas de soporte a la toma de decisiones frente a realidades complejas y
8. La fuerte presión de los productos comerciales disponibles.

#### Aplicaciones de la minería de datos

Una de las tareas más importantes de la minería de datos es la identificación de aplicaciones para las técnicas existentes y desarrollar nuevas técnicas para dominios tradicionales o de nueva aplicación.

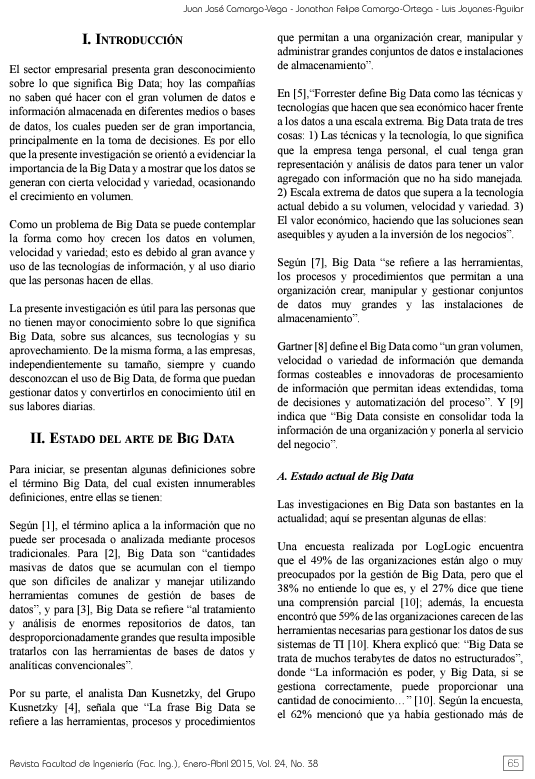
Existen múltiples áreas en donde la minería de datos se puede aplicar, prácticamente en todas las actividades humanas que generen datos, es decir, puede ser aplicado en:

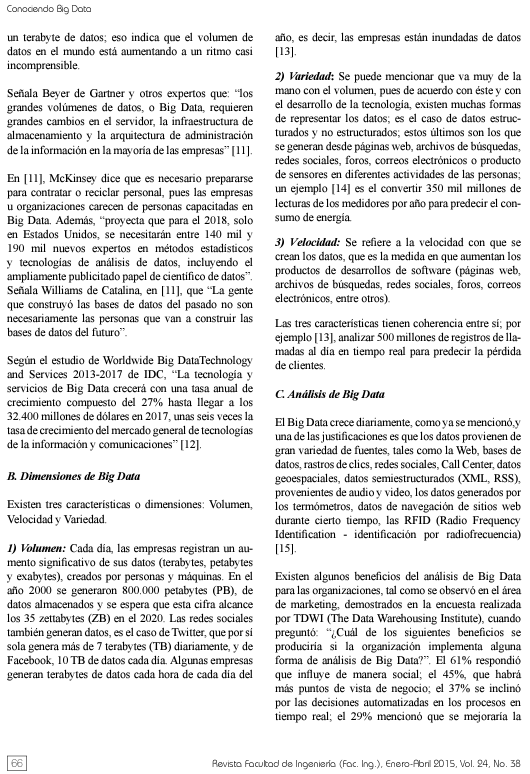
* Comercio
* Medicina y farmacia
* Seguridad y detección de fraude
* Ciencias ambientales
* Ciencias sociales

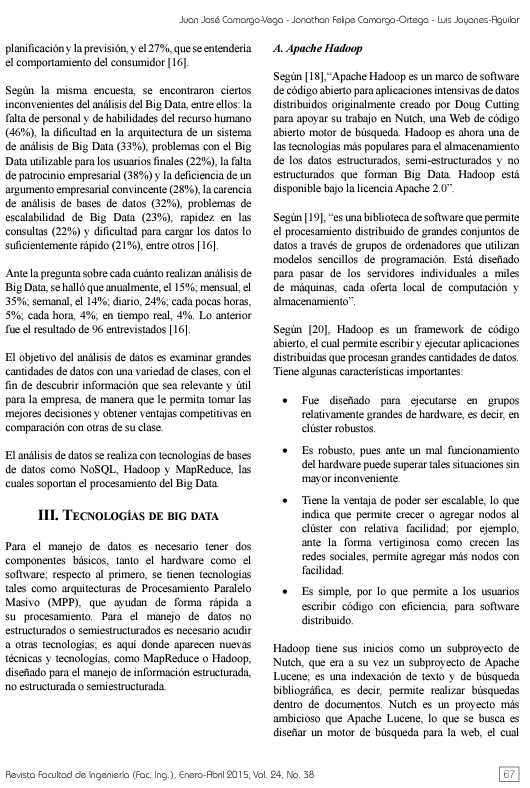
#### Inteligencia artificial

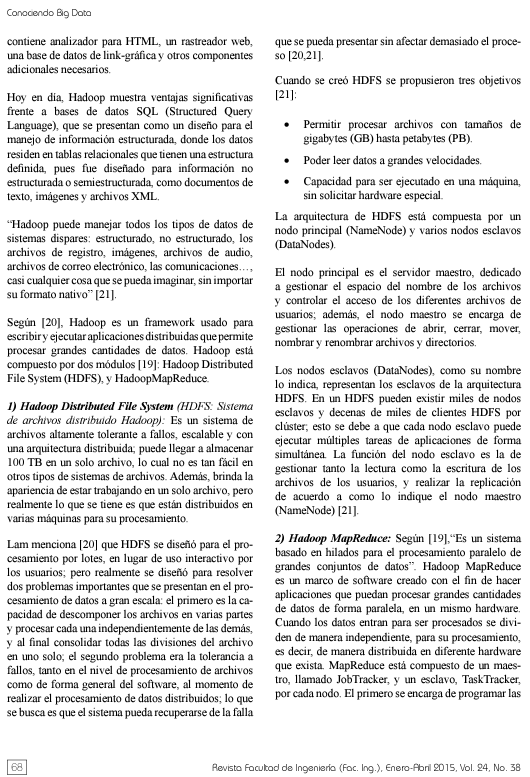
La inteligencia artificial (IA) es la simulación de procesos de inteligencia humana por parte de máquinas, principalmente sistemas informáticos.

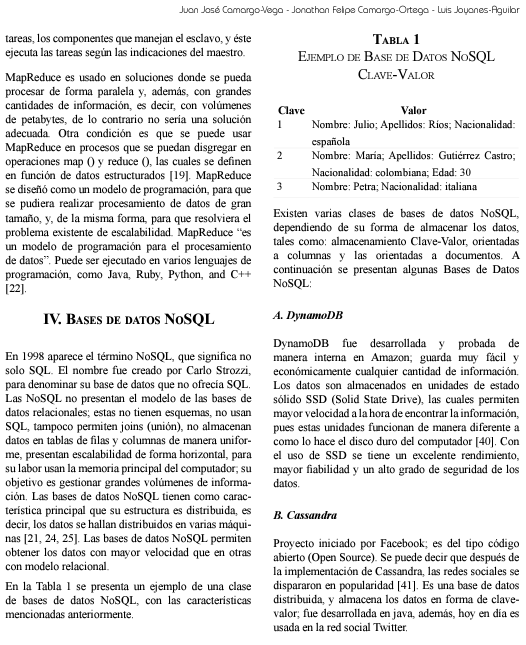
Incluye el aprendizaje, es decir, la adquisición de información y reglas para el uso de la información, así como también el razonamiento, usando las reglas para llegar a conclusiones aproximadas o definitivas y la autocorrección.

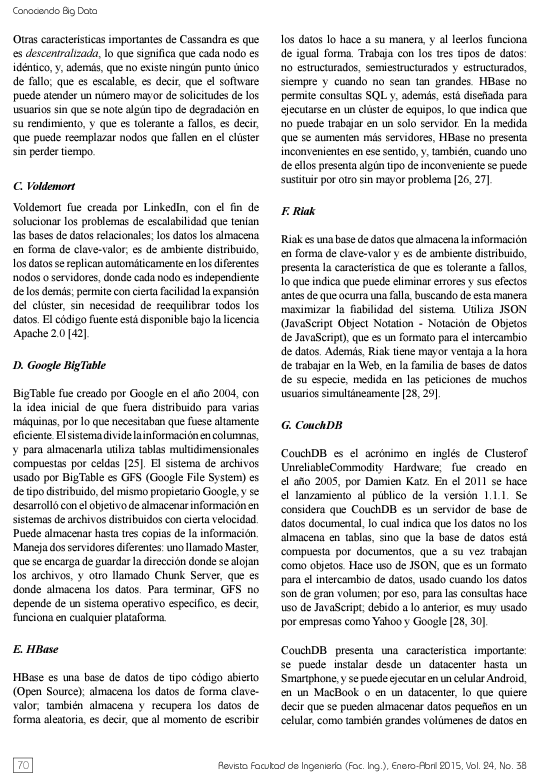


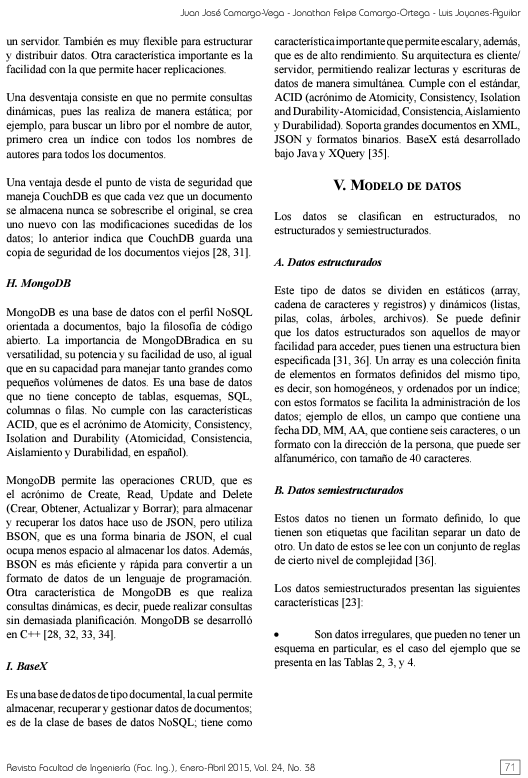


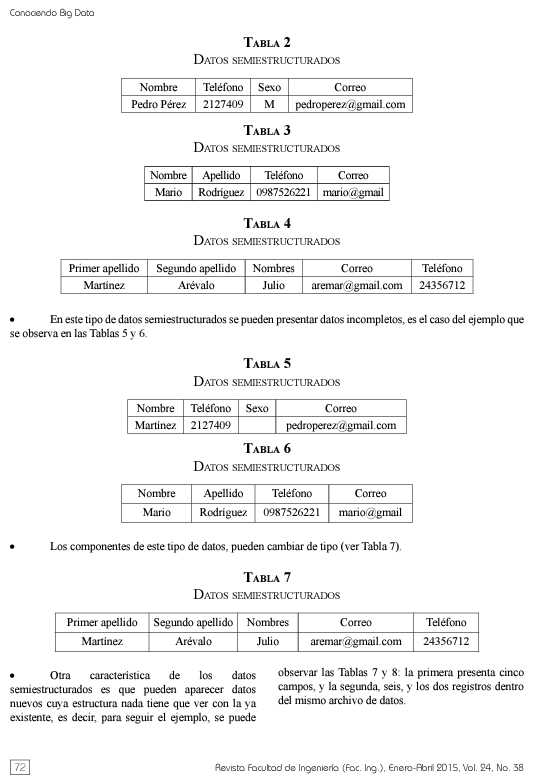


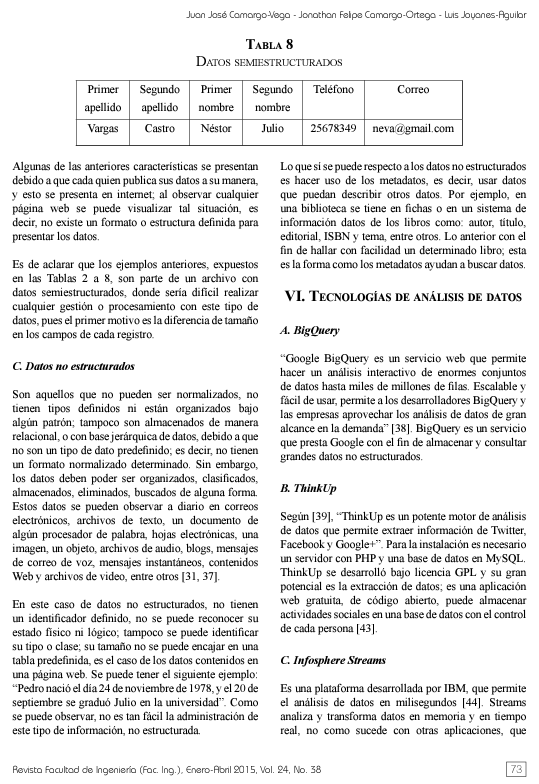


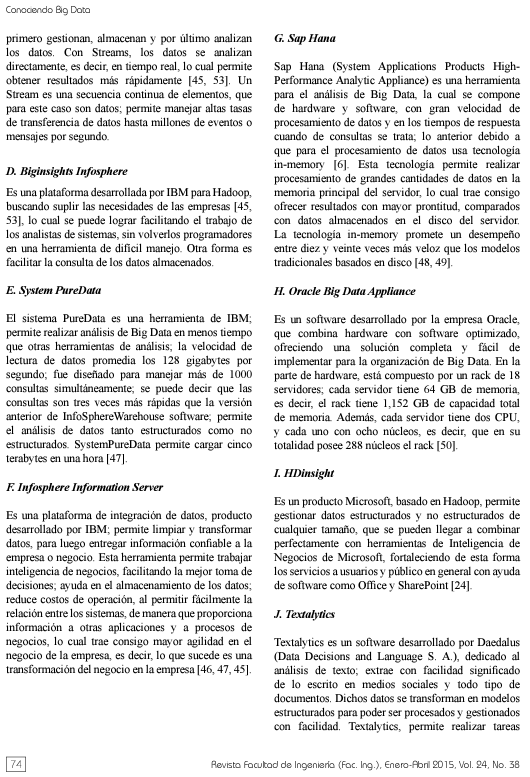


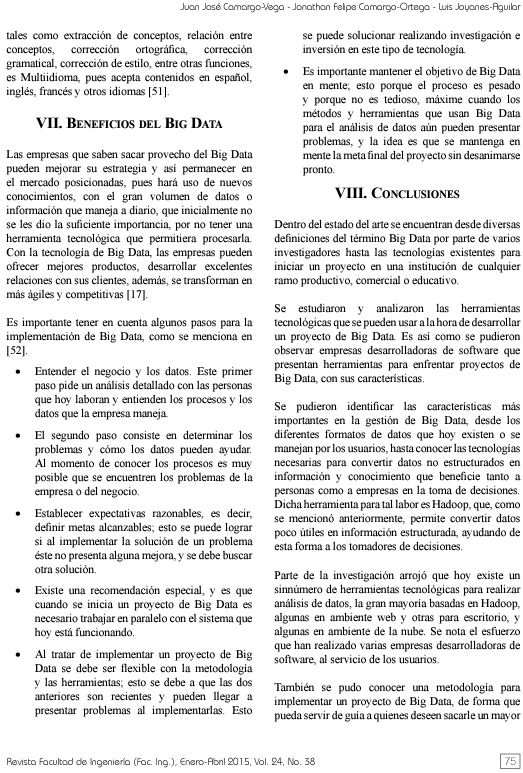




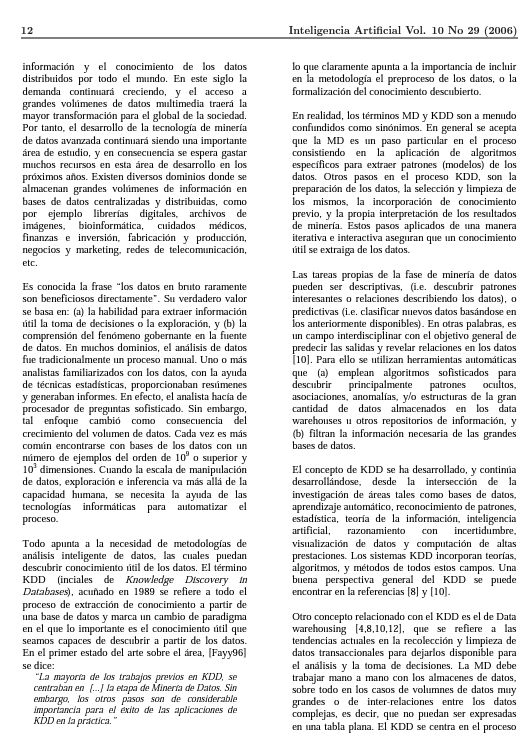


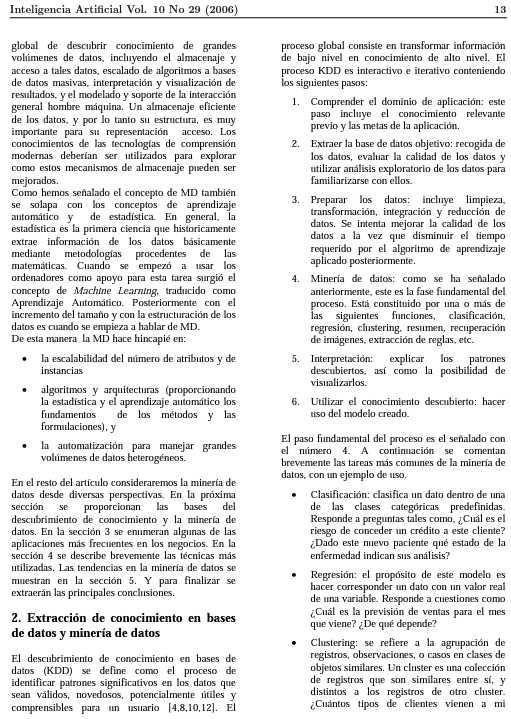


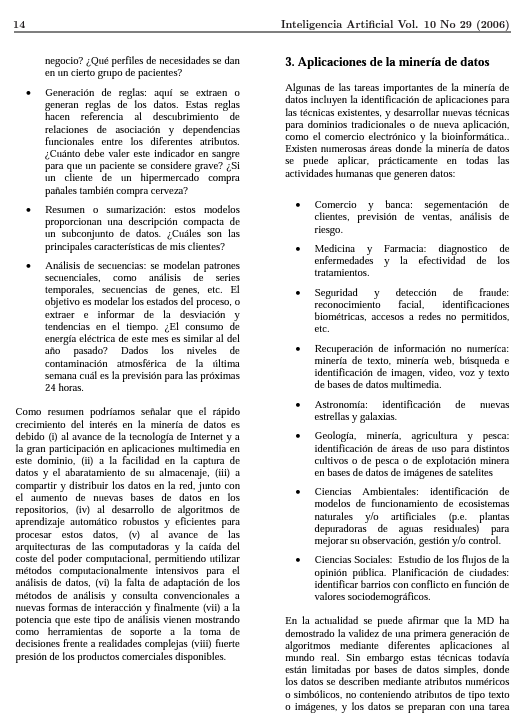


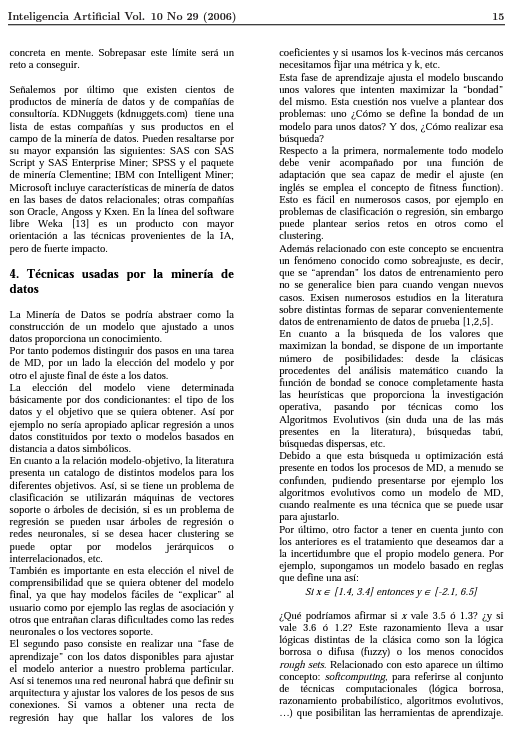
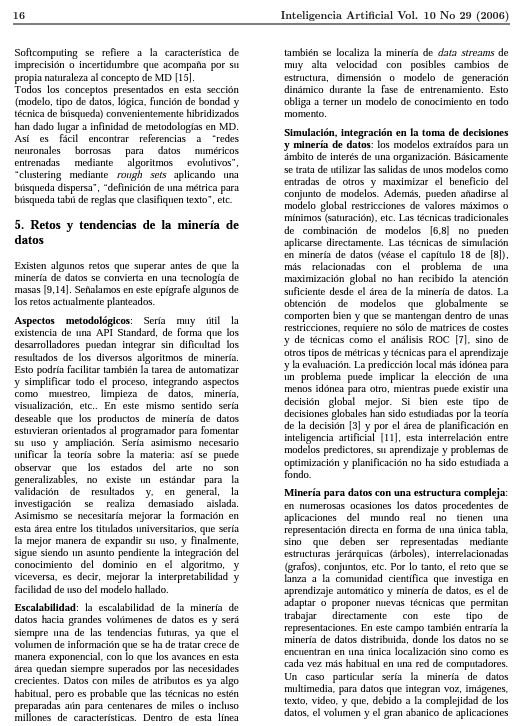


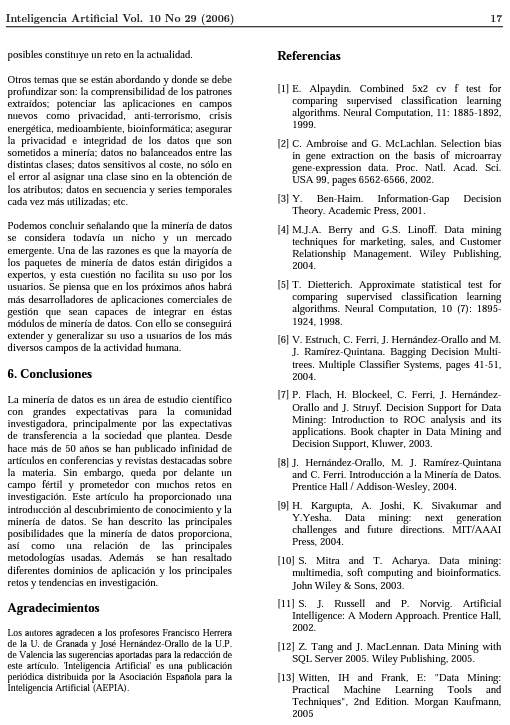










**4.2. Ciencia de datos**

**Introducción:**

A pesar de las grandes inversiones en equipos de ciencia de datos, muchas empresas no visualizan el valor de sus datos. En la carrera por contratar talento y crear sistemas de ciencia de datos, algunas organizaciones experimentan flujos de trabajo ineficientes para el equipo, donde diferentes personas utilizan herramientas y procesos que no funcionan bien sin una administración centralizada más disciplinada, la tendencia es que los ejecutivos no obtengan un retorno completo de sus inversiones.

La mayoría de las organizaciones hacen de la ciencia de datos una prioridad e inversión. Actualmente, las empresas califican la analítica y la inteligencia empresarial como las tecnologías de diferenciación más importantes para sus organizaciones, los directivos ven estas tecnologías como las herramientas más estratégicas para las empresas, por tal motivo están realizando grandes inversiones.

Ya que el acceso a los datos lo debe otorgar un administrador de base de datos, los gestores de datos constantemente tienen que esperar por la liberación de los datos y los recursos que necesitan para su estudio. Una vez que se tiene el acceso, el equipo puede analizar los datos a través de diferentes tecnologías posiblemente incompatibles. Por ejemplo, un científico podría desarrollar un modelo utilizando el lenguaje R, pero la aplicación en la que se usará está escrita en un lenguaje distinto, es por esto que la implementación de los modelos en las aplicaciones útiles puede demorar semanas o incluso meses. Las empresas detectaron que, si no cuentan con una plataforma integrada, el trabajo de la ciencia de datos es ineficiente, inseguro y difícil de escalar. Esto llevó al desarrollo de [plataformas de ciencia de datos](https://blogs.oracle.com/datascience/why-data-science-platform), en las cuales se lleva a cabo todo el trabajo de ciencia de datos.

Una plataforma estable alivia muchos de los desafíos de la implementación de la ciencia de datos, también ayuda a las empresas a convertir sus datos en información de forma más rápida y eficiente. Con una plataforma centralizada de aprendizaje automático, los científicos de datos pueden trabajar en un entorno de colaboración con sus herramientas de código abierto favoritas, donde todo su trabajo es sincronizado mediante un sistema de control de versiones.

**Objetivo:**

Comprender las bases de la ciencia de datos, con el fin de conocer los fundamentos matemáticos, económicos y administrativos para la ciencia de datos en una organización.

Los temas a revisar para responder al objetivo de esta sesión son:

* Ciencia de datos: descripción, funciones y alcance
* Fundamentos matemáticos, económicos y administrativos para la ciencia de datos

**Datos**

Hace referencia a la mínima unidad del valor respecto del análisis que puede tener un sentido completo. Es el equivalente a un punto en nuestro espacio relevante, es también el elemento más crudo del cual disponemos y que será consumido por el proceso de análisis.

En cuanto a sus características, se pueden mencionar:

* **Origen:** fuente de referencia, así como el medio para obtenerlo.
* **Tipo:** semántica del valor representado, sea texto, número, posición, código, etc.

**Evolución en el almacenamiento de datos**

El almacenamiento y recuperación de datos fue una de las primeras aplicaciones llamadas “Cerebros Electrónicos”. La posibilidad de manejar cantidades de información trajo problemas técnicos, por ejemplo, si se deseaba tener el censo ordenado por apellidos, y también por códigos postales, lo que se hacía era duplicar la información en dos cintas, una por cada orden.  
  
El problema surgía cuando, por ejemplo, había que corregir la dirección de un individuo censado, en este caso, podría ocurrir que la persona encargada de la corrección, por descuido, modificara solo una de las cintas.

El resultado sería que, al consultar posteriormente la información, dicho individuo apareciera con la dirección diferente según el método de ordenación escogido según la cinta consultada, por lo tanto, las bases de datos relacionales han sido capaces de gestionar y mantener la integridad de los datos.

**Tecnologías Big Data**

Para dar uso a las tecnologías de la informática es necesario identificar las fases de un proyecto de extracción de conocimiento a partir de datos, para apoyar la toma de decisiones que incluirá la obtención de los mismos, su integración y consolidación en un repositorio, la consecuencia de información mediante métodos analíticos, su presentación y finalmente su uso.

Las tecnologías de Big Data son aquellas que dan soporte a la captura, transformación, procesamiento y análisis de los datos, ya sean estructurados, semi- estructurados o no estructurados.

Las aplicaciones se describen en lenguaje de programación, que disecciona peticiones sobre porciones de datos almacenadas en centenares de ordenadores de forma paralela, luego combina la información de todas ellas para proporcionar la respuesta a la petición. Motores SQL como Dremel (Google), Hive (Hortonworks) y Spark (Databricks) proporcionan tiempos de respuesta muy cortos en los procesos.

**4.3. Ciencia de datos en los indicadores del negocio**

**Introducción:**

Crear, evaluar, implementar y monitorear modelos de aprendizaje automático es un proceso complejo, es por esto que las herramientas de ciencia de datos se incrementaron. Los científicos de datos utilizan muchos tipos de herramientas, una de las más comunes son los cuadernos de código abierto, estos son aplicaciones web para escribir y ejecutar código, visualizar datos y ver resultados, todo dentro de un mismo entorno. Para identificar qué herramienta de ciencia de datos es adecuada, es necesario formular las siguientes preguntas: ¿Qué lenguajes utilizan sus científicos de datos? ¿Qué métodos de trabajo prefieren? ¿Qué fuentes de datos usan?

La ciencia de los datos ejecuta el proceso de convertir los datos en información y ésta en conocimiento, a través de técnicas matemáticas, estadísticas y analíticas computacionales, es por esto que la ciencia de datos es un área multidisciplinaria que persigue el objetivo de mejorar la toma de decisiones.

Datos ⇒ Información ⇒ Conocimiento ⇒ Toma de decisiones

Para lograr su objetivo, utiliza tres principales técnicas: la [minería de datos](https://es.wikipedia.org/wiki/Miner%C3%ADa_de_datos), la [inferencia estadística](https://economipedia.com/definiciones/inferencia-estadistica.html) y el [aprendizaje automático](https://es.wikipedia.org/wiki/Aprendizaje_autom%C3%A1tico).

Básicamente, la importancia de la ciencia de los datos aplicada en la empresa está en que, gracias a ella, podemos desarrollar un conocimiento muy profundo del negocio, un conocimiento que de ninguna otra forma pudiésemos obtener. Es el proceso de exploración y de análisis sobre un conjunto de datos para encontrar patrones dentro de este, dichos patrones han de ser útiles para la toma de decisiones.

Una de las técnicas computacionales mediante la cual se consigue la inteligencia artificial, en el contexto de la ciencia de los datos, es el Machine Learning (ML), este ayuda a que los sistemas de información puedan ejecutar de manera muy precisa todo el procedimiento de ciencia de datos de manera automatizada, solo así es posible que las empresas y organizaciones se beneficien de la inteligencia artificial y sus componentes.

**Objetivo:**

Analizar la relevancia de la ciencia de datos en los indicadores de negocios, con el fin de saber su uso dentro de una organización y conociendo las oportunidades a futuro que pueda haber en esta.

Los temas a revisar para responder al objetivo de esta sesión son:

* Ciencia de datos en los indicadores del negocio

**Aprendizaje automático**

El aprendizaje automático se relaciona con el área de la informática y las ciencias de la computación, esta trata de la arquitectura de modelos y desarrollo de programas informáticos que aprenden a resolver problemas, potencialmente complejos, a partir de ejemplos o datos que se toman como *input* mediante alguna instrucción.

En este contexto, el aprendizaje es la capacidad de programar algoritmos de búsqueda progresivos a la solución de problemas específicos, a medida que se va suministrando un conjunto mayor de datos o ejemplos de partida.

**Usar datos en una empresa**

Resulta necesario que las organizaciones comiencen, culturalmente, a aprender a sacarle el mayor valor posible a los recursos que posee, por lo que es importante mencionar tres aspectos claves: identificación y encuadre de la situación, interacción con perfiles cuantitativos y distintos tipos de análisis; no solo para la analítica, sino para el uso de datos y evidencia en general.

Esto quiere decir que el uso de los datos y de información no está visto como la órbita de un pequeño grupo de elegidos, sino como una actividad común de todos los integrantes de la organización.

**Mirando al futuro**

La principal herramienta para entender las implicaciones sociales de la tecnología es la educación. Los especialistas deben poner en medios para que la sociedad comprenda cómo funcionan los algoritmos en torno a la ciencia de datos. Pero del mismo modo los ciudadanos deben hacer un esfuerzo por entender el mundo digital.

Es difícil cuantificar sus efectos concretos, pero parece evidente que muchos trabajos desaparecerán, especialmente aquellos más técnicos y fácilmente automatizables, dando como consecuencia que otros nuevos sean creados, traduciéndose como una amenaza, pero también una oportunidad.

**Inteligencia de Negocios (BI)**

La Inteligencia de Negocios fue introducida por Howard Dresner en el año 1989. En la actualidad, uno de los campos en los que la BI es más utilizada por sus excelentes resultados es en las finanzas porque permite la visualización, análisis, comprensión y seguimiento de la información en tiempo real, de manera sencilla y efectiva.

“La Inteligencia de Negocios no es solamente un producto, sino una herramienta que utiliza diferentes tecnologías y en ellas asocia y combina métodos efectivos con determinados productos, para organizar conjuntos de datos, cuyo uso e interpretación es relevante para mejorar las utilidades y desempeño de un negocio, además también plantea que tal herramienta permite construir y aplicar mecanismos capaces de acelerar ciertas acciones y disposiciones sobre el funcionamiento de los negocios, así como la sistematización de la información clave para la toma de decisiones acertadas”.

*(Murillo, Cáceres Castellanos, 2013, p.121)*