

UNIDAD 1: FUNDAMENTOS DE LA INTELIGENCIA DE NEGOCIOS Y SU APLICACIÓN EN LA ADMINISTRACIÓN DE EMPRESAS.

Autor:

Ing. César Bustamante Chong, Mgtr.

Docente Tiempo Completo de la Facultad de Ingenierías de la Universidad Tecnológica ECOTEC.

cbustamante@ecotec.edu.ec

1.1 Introducción

En la actualidad, escribir sobre la información que se genera, gestiona y aplica; tomaría mucho tiempo en explicar, puesto que la sociedad desde hace treinta años ha estado expuesta a un crecimiento exponencial del manejo de los datos. En este sentido, existen mediciones y explotación de estos datos en lugares nunca antes experimentados y esta progresión y evolución, con el tiempo se ha visto reflejado en poderosas herramientas de software que permiten realizar cálculos, predicciones, estimaciones con un porcentaje muy alto de credibilidad. Todo esto se realiza con base en datos históricos, que anteriormente se perdían por estar almacenados en papeles o carpetas. Sin embargo, gracias al avance de los dispositivos de almacenamiento se puede preservar un volumen importante y de varios años de antigüedad sobre cualquier información.

Resultado de esta cantidad ingente de información, también se observaron múltiples problemas; entre los cuales se pueden anotar:

- La manera de buscar ciertos datos sin afectar el rendimiento de la consulta,
- La forma de diferenciar los datos encontrados,
- La utilidad del propósito de la búsqueda,
- La aplicación de los datos como conocimiento útil, alineados a los objetivos y metas de la organización.

Como consecuencia de lo expuesto, surgieron un sinnúmero de conceptos aplicados, tales como almacén de datos, minería de datos, máquina de aprendizaje, inteligencia de negocios (entre los más importantes). Todo ello, evolucionó hasta su producto final, con el propósito de profundizar en la información y transformarla en un conocimiento útil, eficiente y eficaz para la toma de decisiones, en el control de procesos y en la predicción de ciertos parámetros empresariales.

1.2 Evolución de la inteligencia de negocios.

En la actualidad, la herramienta de Inteligencia de negocios (IN) conocida también por sus siglas en inglés (BI), ha constituido una solución para la evolución de la gestión empresarial. Sin embargo, el término BI no es reciente, ya que en 1958 se le adjudicó sus inicios al prominente investigador alemán Hans Peter Luhn (quien trabajó para IBM) y creó más de 80 patentes. También, Hans Peter es conocido por ser el primero en aplicar la estadística en los análisis textuales de recuperación de información e inventor del algoritmo de Luhn (algoritmo de módulo 10, que permite detectar errores en la transcripción de dígitos). Asimismo, la publicación de su artículo titulado “A Business Intelligence System”, aun cuando fue una investigación básica, se le consideró un visionario para su época, ya que afirmó que: “BI es la habilidad de aprender las relaciones de hechos presentados de forma que guíen las acciones hacia una meta deseada” (Luhn, 1958, p. 314).

Poco después, el canadiense Kenneth Iverson fue un notable matemático, que desarrolló una nueva notación para operaciones sobre series numéricas. En aquella época, la empresa IBM lo contrató, ya que esta crea un intérprete para ejecutar dicha notación y en 1962 realiza un importante avance para la consecución de lo que ahora es el BI. Por esta razón, publica una descripción de su notación en un lenguaje de programación multidimensional (APL); el cual es el pilar fundamental para el procesamiento analítico en línea, conocido como OLAP.

Siete años después, en 1969, se creó el concepto original de las bases de datos formulado por Peter Chen en su obra “The Entity Relationship Model-Toward a unified view of data”, quien se convirtió en el padre del modelo de entidad relación (MER). Este modelo fue un gran paso y se convirtió en la pieza angular para

diversas metodologías sobre análisis y diseño de sistemas, herramientas de ingeniería de software asistido por computadora y repositorios de datos. Hasta el momento, este modelo ER es uno de los trabajos más citados en las ciencias de la computación.

Posteriormente, en la década de los setenta se crearon los grandes sistemas de gestión de base de datos y las primeras aplicaciones empresariales (SAP, JD Edwards, Siebel, PeopleSoft, entre otros). Estos sistemas de gestión de datos permitían realizar el acceso a la información almacenada (data entry) proceso de ingresar datos mediante una computadora. Entre las más usadas se encuentra una herramienta ofimática (aplicación informática), en cambio, estos no fueron capaces de ofrecer un acceso rápido y fácil, por cuanto su organización era de muy baja calidad, lenta y de difícil acceso a información específica, tal como indica Cubero & Berzal (2011). En esta época aparece la interacción con el usuario mediante la posibilidad de usar el “scroll”; pero aun así los informes, en algunos casos, seguían siendo estáticos.

En la década de los ochenta vio la luz, el concepto de almacén de datos (en inglés Datawarehouse) y según Conesa & Curto (2010) es un repositorio de datos que proporciona una visión global, común e integrada los datos de una organización, con las siguientes propiedades: estabilidad, coherencia, fiabilidad y con información básica. Asimismo, uno de los artífices originales de la creación de los almacenes de datos es Ralph Kimball, quien diseñó una metodología llamada “-modelado dimensional- considerada un estándar en los sistemas de soporte de decisiones, al expresar lo siguiente: es un almacén de datos que extrae, limpia, conforma y entrega una fuente de datos dimensional para la consulta y el análisis” (Kimball & Ross, p. 7). Con lo cual, este autor tiene un punto de coincidencia con Dresner (2010) al determinar que el almacén de datos es la unión de todos los “data-marts” de una entidad.

Para esta época se dio un impulso a los primeros sistemas de reportes, el cual consta de un servidor de informes y una aplicación local para su definición. El primero permite el almacenamiento centralizado de las plantillas y la generación a petición del usuario de un nuevo informe; mientras que el segundo, permite la

edición y creación de las plantillas. También, en septiembre de 1985, lanzan al mercado la hoja de cálculo más popular y utilizada hasta ese momento: Excel 1.0, lo cual constituyó una disyuntiva para esta época; puesto que existían sistemas de gestión con base de datos potentes, pero no existían aplicaciones que pudiesen utilizar su explotación.

No obstante, en la década de los noventa, empezaron a surgir y fueron populares las herramientas de BI; las cuales ofrecían acceso a un sinnúmero de bases de datos y a la información estructurada generada por las empresas. En 1989, un ilustre desconocido llamado Howard Dresner comenzó a acuñar el término de BI, para describir –los conceptos y métodos con el objetivo de mejorar la toma de decisiones empresariales mediante el uso de sistemas basados en hechos de apoyo- y con el pasar del tiempo hasta finalizar la década del 90 se comenzó a popularizar estas herramientas.

Por esta razón, se crearon múltiples aplicaciones de BI que ofrecían acceso a las diferentes bases de datos y a la información estructurada generada por las empresas. Sobre todo permitían tener reportes operacionales, modelación estadística para campañas publicitarias, ambientes OLAP multidimensionales (procesamiento analítico que implica grandes lecturas de datos para llegar a extraer una información útil) para analistas, dashboard (representación gráfica de las principales métricas para conseguir los objetivos de una compañía) y scorecard (una herramienta que relaciona estrategias y objetivos claves de desempeño y resultados en una organización) para ejecutivos, entre otros.

En el nuevo siglo (popularmente llamado Y2K), la evolución de las herramientas de BI cambió drásticamente, puesto que no solamente consideraba a la información estructurada y relacionada con las bases de datos, sino que también comenzó a considerar aquel tipo de información y documentos no estructurados; es decir, imágenes, videos, audios, inclusive correos electrónicos, a esta nueva versión se le denominó BI 2.0. (Jeff, 2017).

De esta forma se crean nuevos conceptos, producto de la enorme cantidad de información que se almacenan en las bases de datos relacionales y no relacionales

(Big Data). En tal sentido, es fundamental la creación de herramientas que estén dispuestas a analizar y categorizar esta información de forma rápida y ágil (minería de datos y minería de textos) mediante resultados que permitan la toma de decisiones o predecir resultados (análisis estadístico y análisis predictivo) que aseguran la ventaja competitiva de una organización.

En los últimos años, gracias a la evolución de los datos estructurados y no-estructurados, se han visto afectadas múltiples disciplinas de las ciencias. La ciencia de la estadística, minería de datos, analítica predictiva y el aprendizaje autónomo propició la aparición de un nuevo campo de la ciencia interdisciplinaria, la cual muchos especialistas han llamado como la ciencia de los datos.

Según Hey, Tansley, & Tolle (2009) en una recopilación de artículos en honor a Jim Gray, ganador del premio Turing; quien imaginó a la ciencia de los datos como un “cuarto paradigma” de la ciencia (empírico, teórico, computacional, basado en datos) y afirmó que todo lo relacionado con la ciencia está cambiando vertiginosamente debido al impacto de la tecnología de la información y la cantidad ingente de datos. Esta nueva ciencia se apoya muchas veces en modelos, ecuaciones, algoritmos y la evaluación e interpretación de los resultados.

El término ciencia de los datos fue acuñado por primera vez en el año 1996, durante una conferencia de los miembros de la IFCS (International Federation of Classification Societies). Poco después en el 2001, Cleveland (2001) manifiesta que la ciencia de los datos es una disciplina que involucra al campo de la estadística incluyendo avances de la computación con los datos. Para ello se establece seis áreas técnicas que conformarían el campo de ciencias de los datos: investigaciones multidisciplinarias, modelo y métodos para datos, computación con datos, pedagogía, evaluación de herramientas y teoría.

Conforme avanza el tiempo, poco a poco ha ido avanzando esta nueva ciencia, hasta el punto que en el 2015 la importante editorial estadounidense Springer (especializada en libros y revistas académicas) lanza una revista llamada “Journal on Data Science and Analytics”, en la cual se pretende publicar trabajos originales en ciencias de datos y analítica de big data.

1.3 Herramientas de inteligencia de negocios.

A finales de la década del 2010, la mayoría de las personas interactúan por medio de un dispositivo electrónico de cualquier índole, lo que conlleva a un alta –y hasta cierto punto incuantificable- cantidad de información. Esta debe ser recuperada, almacenada y muchas veces analizada con fines de obtener un conocimiento pleno para predecir eventos u casos o en otros momentos para buscar patrones o tendencias en la información.

Por esta razón la evolución del BI, surge de una serie de herramientas que son utilizadas principalmente por los analistas de datos para entregar análisis e informes con resultados en consultas de carácter empresarial. Sin embargo, no resulta novedoso que ejecutivos medios de la organización o el mismo personal operativo estén utilizando las herramientas de BI, gracias en parte al alto desarrollo de fácil operación y uso de los mismos como herramientas del descubrimiento de datos.

Estas herramientas de BI, pueden a construir tableros de control sobre indicadores de desempeño en la organización como manifiesta (Bonney & Armijo, 2005), diseñando gráficos y otro tipo de infografías que permiten observar las métricas empresariales y poder tomar decisiones en firme. Por ello, se reconoce que existen múltiples herramientas de inteligencia, que combinan algunas funciones en una sola o se pueden apreciar por separado; sin embargo, entre las principales funciones se pueden encontrar las siguientes:

Minería de datos

Es el conjunto de técnicas y tecnologías de la información que permiten explorar una gran cantidad de datos almacenadas en múltiples bases de datos con la finalidad de detectar información procesable para deducir patrones repetitivos, tendencias o reglas que existen en los datos y expliquen su comportamiento en un determinado contexto. Este proceso se realiza mediante un algoritmo de búsqueda (muy próximo a la determinación de inteligencia artificial y/o redes neuronales), análisis matemático o una tendencia (Nihad A Hassan, 2018).

Estos patrones, tendencias o reglas suelen recopilarse, definirse y aplicarse en un modelo de minería de datos; este modelo es la aplicación de un tratamiento a una cantidad específica de datos para obtener información de ellos. Existen dos modelos considerados los más usados y comunes en el momento de realizar el análisis y son:

- 1) Modelo de red neuronal: es un modelo computacional basado en un gran conjunto de neuronas simples (neuronas artificiales), que se comporta aproximadamente análoga al comportamiento observado en las neuronas de los cerebros biológicos; este modelo trata de múltiples entradas de datos, se agrupan estos datos para luego realizar un análisis basado en un algoritmo específico.
- 2) Modelo de árbol de decisión: es un modelo de predicción basado en un conjunto de datos que se agrupan según características específicas en diagramas de construcciones lógicas, similares a los sistemas de predicción basados en reglas, y muestran una serie de condiciones que ocurren en forma sucesiva para resolver un problema.

Sin embargo, también suelen utilizarse modelos que se basan en:

- Regresión lineal: Es una técnica estadística que estudia la relación entre variables, se utiliza muchas veces para predecir un amplio rango de fenómenos en diferentes áreas del conocimiento, es la técnica más común utilizada para formar relaciones entre datos.
- Modelos estadísticos: Se basa en una ecuación matemática que reproduce los fenómenos que observamos de la forma más exacta posible y se emplea en todos los diseños experimentales de distancia, se deben tener en cuenta los datos ingresados y la influencia del azar y la probabilidad que se tenga en las observaciones.
- Agrupamiento o clustering (algoritmo k-means/algoritmo k-medoids): consiste en dividir la información en varios grupos en la base de datos, el objetivo es encontrar grupos que sean diferentes entre sí, y sus miembros que guarden similitud uno del otro; una persona con gran experiencia y conocimiento en la organización debe identificar los clusters o grupos.

- Reglas de asociación: consiste en la exploración de la información con el objetivo de identificar relaciones o hechos que ocurren entre valores dentro de la base de datos. Según Weiss & Indurkha (1997) dependiendo del objetivo del análisis de los datos, los algoritmos se pueden clasificar en supervisados o predictivos (predicen un dato) y no supervisados o del descubrimiento del conocimiento (descubren patrones y/o tendencias de datos).

Estos modelos de minería de datos se pueden aplicar en varios escenarios, tales como:

- Pronóstico: cálculo y predicción de datos específicos (cálculo de ventas, predicción de unidades producidas).
- Riesgo y probabilidad: selección y asignación de probabilidades, determinación de puntos de equilibrio en escenarios de riesgo (selección de escenarios posibles para minimizar riesgos en fondos de inversión).
- Recomendaciones: recomendaciones de seleccionar dos o más datos específicos (determinar las promociones de productos, generar recomendaciones).
- Búsqueda de secuencias: análisis de los datos específicos para predecir posibles eventos (análisis de los artículos, qué clientes buscan y predecir futuras ventas).
- Agrupación: análisis y predicción de afinidades para datos específicos.

En (Microsoft, 2018) y luego (Han, Kamber, & Pei, 2012) muestran los pasos que se siguen para definir y construir un modelo de minería de datos:

Definir el problema, que consiste en definir exhaustivamente el problema y considerar las formas de usar los datos para resolver el mismo; y se deberán contestar las siguientes preguntas:

- ¿Qué está buscando? ¿Qué tipos de relaciones intenta buscar?

- ¿Refleja el problema que está intentando resolver las directivas o procesos de la empresa?
- ¿Desea realizar predicciones a partir del modelo de minería de datos o solamente buscar asociaciones y patrones interesantes?
- ¿Qué resultado o atributo desea predecir?
- ¿Qué tipo de datos tiene y qué tipo de información hay en cada columna? En caso de que haya varias tablas, ¿cómo se relacionan? ¿Necesita limpiar, agregar o procesar los datos antes de poder usarlos?
- ¿Cómo se distribuyen los datos? ¿Los datos son estacionales? ¿Los datos representan con precisión los procesos de la empresa?

Al responder estas preguntas, se debe realizar un estudio de disponibilidad de datos para verificar las necesidades de los usuarios con respecto a si se encuentran disponibles o no los datos en la organización.

Preparar los datos, este paso consiste en consolidar y limpiar los datos que se encuentran en la organización. Muchas veces los datos suelen estar dispersos, almacenados en diferentes repositorios de datos y en algunas ocasiones en diferentes sitios físicos; así también se espera obtener datos con ciertas incoherencias en cuanto a las entradas o estar repetidas, o inclusive pueden faltar ciertas entradas. Entonces la limpieza de datos no solamente implica arreglar los datos, eliminar los datos inconsistentes, sino que también en muchos casos interpolar valores que faltan, identificar orígenes de datos que puedan faltar (Siegel Clifford, 2014).

Es importante recordar que en esta etapa del modelado se tiene una gran cantidad de datos, por lo cual muchas veces resulta imprescindible trabajar con herramientas de generación de perfiles de datos, de limpieza y/o filtrado automático de datos.

Explorar los datos, este paso consiste en conocer los datos para poder tomar decisiones adecuadas, muchas veces se utilizan técnicas estadísticas que permitan incluir valores máximos y mínimos, calcular la media (revisar máximos, mínimos y valores medios para saber si existen datos representativos de una característica en particular) y desviación estándar (establecen información sobre estabilidad y exactitud de los resultados) y examinar la distribución de datos. Al momento de explorar los datos para conocer el problema empresarial o institucional, se puede decidir si el conjunto de datos contiene datos defectuosos o anómalos, con lo cual se puede esbozar una estrategia para corregir los problemas o se puede obtener una descripción más profunda de los comportamientos o políticas de la organización.

Generar modelos, en este paso se definen las columnas de datos que se usan, se crea la estructura de minería de datos y esta se relaciona con la fuente de datos; en esta fase se generan los agregados y otra información estadística para usar en el análisis; en este instante también se aplican sobre la estructura de minería de datos un algoritmo matemático que pueda indicar los patrones que existen sobre la información, también en algunos casos se ajustan los parámetros en cada algoritmo y se aplica filtros a los datos para crear subconjuntos de datos, pudiendo crear resultados diferentes, pudiendo establecer escenarios diferentes para el análisis y futuras predicciones.

Explorar y validar los modelos, en esta fase se generan los agregados y otra información estadística para usar en el análisis; en este instante también se aplican sobre la estructura de minería de datos un algoritmo matemático que pueda indicar los patrones que existen sobre la información, también en algunos casos se ajustan los parámetros en cada algoritmo y se aplican filtros a los datos para crear subconjuntos de datos, pudiendo crear resultados diferentes, se establecen escenarios diferentes para el análisis y futuras predicciones.

Si ninguno de los modelos que se ha creado puede dar una solución correcta, se deberá volver a generar los pasos anteriores, es decir, volver a definir el problema o investigar los datos originales.

Análisis predictivo:

Es el conjunto de diversas técnicas de la minería de datos y parte de la analítica avanzada que es utilizada para realizar predicciones sobre sucesos futuros desconocidos, que se fundamenta en la extracción del conocimiento mediante información histórica que permita predecir tendencias y patrones de comportamiento en los datos, es muy común tener una circunstancia desconocida de interés que se pretende predecir sus tendencias en el presente o en tiempo futuro, suele también utilizar la aplicación de una técnica estadística o de aprendizaje automático (es un tipo de inteligencia artificial que proporciona a las computadoras la capacidad de aprender, sin programarlas) para crear una predicción cuantitativa sobre el futuro (Siegel, 2013).

La base del análisis predictivo es reconocer las relaciones entre las variables explicativas (también conocida como variable dependiente) y las variables predictivas del pasado (otra variable del experimento que afecta la respuesta y que es conocida como variable independiente) de forma que se pueda adelantar a lo que está por ocurrir. Los datos históricos y transaccionales de la organización se usan muchas veces para identificar los posibles riesgos y oportunidades en el futuro.

El análisis predictivo en algunos casos permite anticipar resultados y comportamientos siendo más detallista, produciendo probabilidades para cada dato buscado, eso lo diferencia de la anticipación; se puede enunciar también, unos modelos predictivos que muestran la relación existente entre el rendimiento específico de un sujeto en una muestra y el análisis de uno o más atributos, características del mismo sujeto. El objetivo de estos modelos predictivos es evaluar la probabilidad de que este sujeto evaluado tenga el mismo rendimiento en otra muestra diferente, en muchos casos y áreas del conocimiento se trata de buscar patrones de datos ocultos que respondan sobre el comportamiento de la variable a investigar.

Estos modelos de análisis predictivo, tienen muchas aplicaciones de las cuales se pueden nombrar:

- En la fabricación de automóviles autónomos, el análisis predictivo sirve para analizar los datos de los sensores sensibles conectados y crean algoritmos para la conducción autónoma de hecho, no solo se aplica en automóviles, sino que sirve en camiones de gran envergadura.
- En la aeronáutica, con el objetivo de reducir costos de mantenimiento y aumentar el tiempo productivo de las aeronaves, se crean aplicaciones de análisis en tiempo real para poder predecir el rendimiento de los subsistemas de navegación en relación con los insumos utilizados (aceite, combustible, despegue, peso de carga, entre otros).
- En servicios financieros, en la predicción de riesgos crediticios, las instituciones utilizan técnicas de aprendizaje automático y herramientas cuantitativas para predecir y mitigar el riesgo.
- En dispositivos en el área médica, permite registrar y analizar ciertos datos en enfermedades de diabetes (por ejemplo) para proporcionar información inmediata con el fin de controlar la enfermedad.
- En aplicaciones de redes sociales, como Facebook se utiliza para personalizar el gusto de cada miembro, es decir, si un miembro se detiene para leer o “dar gusto” en un amigo, pues entonces la red social empezará a mostrar actividad de ese amigo.

El flujo de trabajo para realizar modelos de análisis predictivo, en muchas ocasiones es:

Acceder, recoger y explorar los datos: esta selección de la información se puede dar desde diferentes plataformas: archivos de la organización, base de datos y sensores de los diferentes dispositivos a utilizar.

Procesar y analizar los datos: es el proceso analizar, limpiar, transformar y clasificar los datos con la finalidad de encontrar información útil (patrones) que permitan llegar a una conclusión; también se plantean las características de la extracción, así como trabajar con “datos sucios”

Desarrollar modelos predictivos: es la acción de crear los modelos predictivos específicos sobre el futuro, optimizar los parámetros a analizar, y posteriormente se permitirá la validación del modelo utilizando el análisis de técnicas estadísticas que permiten validar la hipótesis y probarlas utilizando modelos estadísticos estándar.

Integrar el modelo predictivo en un entorno de producción: una vez dispuesto el modelo predictivo que realiza el pronóstico con precisión, se debe integrar a los sistemas de información de la organización, para que comience a poner los análisis a los usuarios mediante páginas web o dispositivos móviles.

Minería de texto

Es el conjunto de técnicas y/o modelo específica de minería de datos y textos analíticos (los documentos y textos de las organizaciones), en unión con herramientas estadísticas que permiten construir una inteligencia predictiva descubriendo tendencias y relaciones en un conjunto de datos estructurados (las bases de datos relacionales) y no estructurados (los datos que provienen de las redes sociales u otros tipos de textos, incluso elementos que se pueden derivar de sus contenidos, como el sentimiento) (Steven, 2018).

La minería de texto se apoya en otras técnicas como la categorización de texto, procesamiento de lenguaje natural, extracción y recuperación de la información, aprendizaje automático, entre otros. Se tiende a hablar sobre “minería de datos y textos” como si fueran lo mismo; sin embargo, cuando se habla estrictamente de minería de textos, se refiere a una forma específica de minería de datos, relacionado exclusivamente con textos.

Las herramientas de minería de texto, tienen muchas aplicaciones porque es un método muy eficiente para generar nueva información y conocimiento, tales como:

- En los sitios web que manejan correos electrónicos, permitiendo realizar un filtrado más eficiente y seguro para correos spam; en redes sociales para analizar datos de gusto o relevancia de noticias; en determinar

opiniones de usuarios sobre temas particulares entre clientes y ciertos productos.

- En la seguridad de entidades financieras, para prevenir fraudes y delitos cometidos en internet; contra la lucha del terrorismo mediante el análisis de blogs y otras fuentes en línea permitiendo advertir sobre posibles amenazas.
- En la investigación, en ciencias como la atención médica y el diagnóstico sanitario permitiendo obtener información relevante en instantes que permiten tomar decisiones acertadas y en forma ágil.

Esta técnica es relativamente nueva, pero en términos generales la extracción de textos se puede decir que se desarrolla en estas etapas:

Determinación de los objetivos y recolección de los datos: se define qué es lo que se busca en la investigación, limitando hasta los puntos que se quiera profundizar y el alcance que pueda tener, definiendo claramente los límites; adicionalmente se recolecta los datos de diferentes fuentes de información (tales como sitios web, correos electrónicos, comentarios de clientes, archivos de documentos, etc.).

Pre-procesamiento de los datos: consiste en la definición, selección, análisis y clasificación de los textos o documentos mediante el procesamiento del lenguaje natural (hechos, acontecimientos, datos relevantes, relaciones entre ellos, etc.) en la que se extraerá la información en algún tipo de representación estructurada o semi-estructurada que facilita el análisis. (Isson & Harriot , p. 12)

En esta primera parte se debe definir los documentos en forma representativa y debe seleccionarse en forma aleatoria o mediante un muestreo de tipo probabilístico (se debe evitar la duplicación de documentos en esta primera parte). En una segunda parte una vez seleccionado los documentos seleccionados se los tiene que transformar a un formato analizable para poder realizar la representación estructurada o semi-estructurada.

Finalmente, mediante esta estructura se puede reconocer “tokens” (unidades gramaticales más pequeñas, el cual es una cadena de caracteres que tiene un significado coherente en algún lenguaje de programación) lo que implica representar el texto en una lista de palabras mediante otras estructuras (comúnmente árboles), que son más útiles para un posterior análisis, este proceso divide el texto en entidades significativas (palabras, oraciones, etc.) dados los espacios en blanco presentes y puntuaciones inclusive.

Determinación del modelo y análisis de los resultados: después de utilizarse las técnicas estadísticas y disponer del modelo adecuado para conseguir los objetivos planteados; se tratará de determinar a partir de los datos extraídos y analizados, su coherencia y buscar evidencias, similitudes, excepciones que puedan servir al usuario o especialista para llegar a conclusiones en la observación o exploración de los resultados.

1.4 Aplicación de la inteligencia de negocios en la administración de empresas.

A través de los años, uno de los problemas más grandes que enfrentan los altos directivos de una organización es la toma de decisiones. Dependiendo de la forma correcta y muchas veces, por la experiencia acumulada del directivo de la organización, estas decisiones adoptan una ventaja competitiva importante en el mercado. Sin embargo, cuando ocurre lo contrario, dichas decisiones son desastrosa puesto que conlleva a una serie de procesos que pueden conducir al fracaso de la empresa.

Esta disyuntiva para los altos directivos siempre ha causado grandes problemas y su impacto sobre la organización aún más. Pero, en los últimos tiempos las organizaciones se han asociado a las tecnologías de la información. En cada salto tecnológico las empresas siempre han sabido adaptarse a estos cambios y saben sacar su mayor provecho en la innovación actual.

En un estudio realizado por Gartner (2017), las empresas alrededor del mundo están apostando por la analítica del negocio, que será una tecnología primordial para la aceleración en el crecimiento de los mercados. Se estima que el sector de la analítica y la inteligencia de negocios crecerá por encima del 7% en 2017 y reportará un volumen de ventas cercano a los 18.300 millones de dólares. Por otro lado, se estima que el negocio crecerá en los próximos tres años en un 25% debido a que estos sistemas darán accesibilidad, agilidad y conocimiento en los negocios.

En este mismo informe se dice que el mercado correspondiente a las herramientas de analítica de datos e inteligencia de negocios (BI) continuará su avance más rápido que cualquier otro ecosistema de TI. De esta manera, se traslada este mercado a un ecosistema de plataformas de reporting (gestión de informes que facilitan la distribución a diferentes niveles) orientadas al negocio en conjunto con cloud computing (conjunto de servicios que se presentan a través de servidores en internet) y al empleo de analíticas ágiles que incluyen el autoservicio.

En países vecinos como el caso de Colombia, según un estudio de la Universidad Nacional de Colombia (UNAL) del 2017, las herramientas de BI crecen de forma anual en el orden del 11,5% y los gastos de inversión equivalen a cerca de los 7 billones de dólares. De esta manera se demuestra el aspecto trascendental de la aplicación de esta tecnología en este mismo estudio de las 98 empresas que participaron donde el 70% está satisfecho con los resultados del BI.

En el caso de Perú, existe un estudio de López & Guerrero (2017) quienes manifiestan que las Pymes representan el 46% del PIB de este país, lo que resalta la importancia de este grupo en esta región, con su principal reto que es la necesidad de la información para la toma de decisiones por la generación del alto volumen de datos. La inteligencia de negocios y analítica (IN&A) en unión con la tecnología de Cloud Computing (CC) y su modelo orientado a servicios permite acceder a diseñar una solución con bajo costo, con personal y tiempo reducido; la implementación de este modelo permitirá analizar los datos que generan sus operaciones diarias para poder tomar una decisión de negocio fundamentada.

En Ecuador, estas herramientas de BI han encontrado un sinnúmero de aplicaciones en diferentes sectores productivos e inclusive gubernamentales; pero todos apuntan a tener estas cualidades:

- a) Las herramientas deben soportar mayores índices de accesibilidad, agilidad y profundidad analítica.
- b) En las empresas establecidas se tendrá la necesidad de que nuevos desarrolladores conformen los requerimientos, lo cual traerá nuevas disrupciones en el mercado laboral.
- c) En la parte técnica se necesitará un conjunto de datos (datasets) más complejos, lo cual dirigirá la atención de la inversión a lo relacionado con la preparación de los datos.
- d) Con la aparición de datos no estructurados, se apreciará la posibilidad de insertar y extender contenidos como valor agregado en las analíticas.
- e) El uso de datos que proceden de información streaming (descargar y/o reproducción de archivos de audio y video a través de la red de datos) y otros tipos de fuentes se incrementará el uso en las herramientas.
- f) Con el fin de reducir costos y tiempos de integración, la simbiosis entre el cloud computing con la BI se incrementará con el paso del tiempo.
- g) El concepto de marketplaces (Es un sitio donde compradores y vendedores permiten relacionarse y efectuar en una transacción comercial) permitirá a las organizaciones comprar y vender capacidades de analítica e BI.

El ritmo y la revolución vertiginosa de las herramientas de inteligencia de negocios aplicadas en los medios organizativos tanto en la parte privada como en la pública, ha puesto de manifiesto que se apliquen en forma general en las siguientes áreas:

El trabajo del analista

El trabajo del analista de datos mejora mediante el aprendizaje automático y lo primero que se viene a la mente es el significado del aprendizaje automático (“machine learning”), que según (Kaplan, 2017) nos manifiesta que es una rama de la inteligencia artificial cuyo objetivo es desarrollar técnicas que permitan a las

maquinas aprender; es decir, se trata de crear algoritmos capaces de generalizar comportamientos y reconocer patrones a partir de la información suministrada en forma de ejemplos. Por lo tanto, es un proceso de inducción del conocimiento, un método que permite la generalización de un enunciado general a partir de enunciados que describen casos particulares.

Este aprendizaje automático en mucho de los casos se dice que promueve el rápido desplazamiento de los analistas en muchos campos; sin embargo, con el avanzar de la investigación y la tecnología, más bien es todo lo contrario pues se convierte en el mejor asistente que un analista pueda encontrar. Este aprendizaje automático se dice que ayuda al analista en dos factores principales: el analista no pierde tiempo en hacer cálculos matemáticos básicos y puede también explorar y mantenerse en el flujo del análisis de datos, ya que no se detiene y/o pierde tiempo en realizar cálculos.

El resultado que produce el aprendizaje automático, es el cual el analista debe comprender para poder actuar en el contexto (los datos que se introducen no siempre producen resultados con un significado claro); por esta razón, el aprendizaje no debe aplicarse de manera aislada.

Como ejemplo de puede enunciar el siguiente: el impacto en la modificación de un precio en cierto producto; para ello se podría ejecutar una regresión lineal con los datos, que anteriormente se lo hacía en forma manual, después con Excel, R, QView o Tableau; actualmente se realiza mediante el aprendizaje automático y puede visualizarse el consumo en cuestión de minutos, o inclusive segundos y el analista puede pasar a la siguiente fase -analizar en qué parte del tiempo fue un resultado surgido de un factor extrínseco o anormal-, puede ser que haya ocurrido algún lanzamiento de producto por parte de la competencia o existieron factores en la comunicación que influyeron en las compras o el conocimiento del producto.

Análisis de los datos.

En el sector del análisis de los datos, el impacto humano es grande, como anteriormente se ha manifestado que existe una nueva función dentro de las

organizaciones, denominada Analista de datos, el cual posee conocimientos de la Ciencia de los Datos – “Data Science”. Las organizaciones actualmente buscan personas capacitadas para trabajar con datos, y fortalecer sus equipos de análisis. Estas personas que analizan los datos, no son necesariamente personas con alto conocimiento técnico sino la importancia radica en contratar a personas que son capaces de usar los datos y la información para promover cambios y propiciar transformaciones mediante la persuasión.

En tal sentido, a medida que se simplifica el uso de las plataformas tecnológicas se disminuye el interés en especialidades técnicas, el significado se concentra en las personas que centran sus conocimientos en las humanidades, conocimientos más amplios, estas personas posibilitan que exista una ventaja competitiva en las organizaciones. Muchas de estas personas especialistas en humanidades, se pueden observar en empresas como LinkedIn, Paypal, Facebook, Pinterest, entre otras.

Tableros de Control.

La creación de DashBoard o tableros de control con los datos, ha venido evolucionando en gran medida en los últimos tiempos. En términos de diseño “dash”, el cual es un tablero de instrumentos localizado debajo del parabrisas en los automóviles cuyo propósito es tener en forma gráfica los principales indicadores del funcionamiento del mismo. Con esta analogía se tiene que el término “dashboard” se lo aplica en informática para obtener –también en forma gráfica- una serie de indicadores que intervienen en la consecución de los objetivos del negocio y está orientado a realizar cierto análisis para la toma de decisiones y optimizar la estrategia de una empresa; sin embargo, crear el “dashboard” y realizar análisis requiere de cierta habilidad por parte del analista, consiste en contar una historia con los datos.

En situaciones en donde coexisten problemas para lograr una interacción adecuada entre las máquinas y las personas, el análisis pasó de ser puramente científico para hacerse más humanístico. El enfoque transitó de la comunicación de datos a la narración de historias basadas en ellos para así facilitar la toma de decisiones. Un

ejemplo sobre lo antes mencionado es la interesante historia de Nissan y su investigación acerca de los carros autónomos (conducción programada). Esta compañía automotriz, contrató a una doctora en antropología Melissa Cefkin, quien estudió la interacción entre los automóviles que se conducen solos y los seres humanos, lo cual propició un significativo avance; no obstante, aún existen dificultades en entornos mixtos, donde los seres humanos cohabitan con las máquinas.

Procesamiento de lenguaje natural.

El procesamiento del lenguaje natural, constituye una herramienta para el analista invaluable, cuya abreviatura es PLN o NLP (Natural Language Processing) que permite estudiar las interacciones entre las computadoras y el lenguaje humano. Es decir, se ocupa de la formulación e investigación de mecanismos eficaces para la comunicación entre las personas y máquinas por medio de lenguajes naturales -es el idioma o lenguaje hablado o escrito por humanos para propósitos generales de comunicación-, como plantea (Indurkha & Damerou, p. 6).

En otras palabras, existen grandes expectativas alrededor del hecho de que las personas puedan hablar con el software y que este les comprenda, debido a la popularidad que han alcanzado programas como Amazon Alexa, Google Home o Microsoft Cortana. En el caso de este último, al darle la orden de reproducir al clásico “My Way” de Sinatra, los invitados que departen en una reunión pueden escucharla de inmediato.

Dicho concepto puede ser trasladado y aplicado con los datos, es decir, cualquier persona en un lenguaje natural puede hacer preguntas y análisis de los datos con gran facilidad. Un ejemplo interesante sobre dicha aplicabilidad es cuando el alto ejecutivo se encuentra fuera de la oficina, puede utilizar su dispositivo móvil y pedir una información rápida como el total de ventas por cliente en el sector agroindustrial que compró en cuanto a cualquier producto en la región costa, después puede filtrar los resultados, como las ventas realizadas en los últimos 15 días y después agrupar por las diferentes oficinas de esa región. Según Gartner (2017) en el año 2020, el 50% de las consultas analíticas se generarán mediante búsqueda, procesamiento

del lenguaje natural o voz, el cual permitirá a las personas hacer distintos tipos de preguntas sobre los datos y recibir respuestas mucho más relevantes, obteniendo una mejor información para tomar decisiones mucho más acertadas.

Director de datos.

Una nueva función surge con la Tecnología de la Información (TI), en cuanto al director de datos. Desde hace mucho tiempo, en las organizaciones han existido muchos cambios en la gobernanza de TI, desde el jefe de centro de cómputo pasando por el jefe de Sistemas, hasta el actual director de tecnologías de la información, como se le conoce por sus siglas en inglés (CIO). Lo antes mencionado invita a pensar en la idea de que las instituciones evolucionan y priorizan una mayor concentración y responsabilidad en relación con el análisis de los datos.

Este CIO históricamente es también el responsable de la mayoría de las tareas de inteligencia de datos; o sea, supervisa la estandarización, la consolidación y la gobernanza de los activos de datos de toda la organización. Sin embargo, también es responsable de otras iniciativas estratégicas como la arquitectura de TI, la seguridad de los sistemas y las estrategias en las redes empresariales. Aunque en algún momento se generan dificultades entre el CIO y la empresa debido a la sincronía en la velocidad en que se obtiene la información, su seguridad y la gobernanza de los datos; por lo cual, se reconoce la necesidad de tener un responsable de obtener los datos gracias a inversiones en análisis.

Por este motivo, muchas empresas están incorporando en sus filas el perfil del director de datos “Chief Data Officer” (CDO) cuya principal función será tomar decisiones basadas en hechos y evidencias; o sea, se concibe como la persona que analiza, interpreta, clasifica y convierte volúmenes de datos en estadísticas precisas para poder establecer modelos predictivos y prescriptivos, que ayudan a la toma de decisiones. En cambio, en otros casos puede también modificar los procesos empresariales, superar los obstáculos culturales y comunicar el valor del análisis en la organización.

En el pasado el rol del director de datos consistía en realizar tareas de gobernanza y propiedad del software; es decir, asegurarse de que estos están siendo usados por regulaciones relevantes y mejores prácticas, así como seleccionar la gestión de datos al implementar un software analítico y gestionar actualizaciones. En el presente se involucra aún más en la cultura orientada a datos, al desarrollo de estrategias políticas y prácticas para usar los datos como activo corporativo, adopta el rol de liderar nuevas iniciativas de analítica que mejoren las ganancias o bajen los costos. En el futuro el rol del CDO tendría la responsabilidad de la inteligencia artificial, la cual depende de grandes cantidades de datos de entrenamiento y aplicaciones sofisticadas, que los CDO tienen experiencia en gestionar, en el internet de las cosas (IoT) para administrar y utilizar grandes volúmenes de datos (Ed, 2018).

Con esta separación de funciones el CIO puede concentrar su atención en otras funciones, como la seguridad de los datos.

1.5 Glosario de Términos

Inteligencia de Negocios: Es el grupo de herramientas, estrategias, aplicaciones, datos, tecnologías que facilitan la obtención rápida y eficaz de los datos provenientes de diversas fuentes para su posterior análisis e interpretación, brindando este conocimiento para la correcta toma de decisiones empresariales.

Big Data: Es un vocablo acuñado para definir un gran volumen de datos, que pueden ser estructurados o no estructurados que tienen la particularidad de poder ser extraídos, explotados con métodos convencionales de minería de datos, para darles un valor en la toma de decisiones de una empresa.

Aplicaciones empresariales: Es una aplicación informática que está diseñada para satisfacer las necesidades empresariales, para suplir las necesidades personales. En algunas ocasiones se le suele llamar por sus siglas del inglés (E.R.P) que significan “Enterprise Resources Planning”. Esta aplicación permite la integración de ciertas operaciones y áreas de una empresa, además de funcionar como un conjunto de sistemas de información integrado.

Dashboard: Son métricas empresariales representadas en forma gráfica, inicialmente se transformó los datos en conocimiento del negocio, los mismos que intervienen en la consecución de los objetivos de la empresa y está orientado a la toma de decisiones para optimizar las estrategias de negocios, mayor táctica y estrategia.

OLAP: Es un procedimiento informático que es utilizada en el campo de la inteligencia de negocios y sirve para modificar y agilizar las consultas de grandes cantidades de datos, con base en el comportamiento de las reglas del negocio y su proyección en el tiempo. Muchas veces es un análisis dimensional y dinámico de los datos en forma de resumen sobre una empresa.

IOT: Es una definición relativamente nueva, es el acrónimo de las palabras en inglés “Internet of Things” y se refiere a la interconexión actual de todos los objetos de uso común con Internet, se puede decir que es un sistema interconectado de dispositivos electrónicos y eléctricos con identificadores únicos, con la capacidad de transferir datos a través de la red.

CIO: Es la persona que tiene como cargo ser el director de la Tecnología de la Información, como se le conoce en inglés “Chief Information Officer” (CIO); es aquella persona responsable del desarrollo, la implementación y la operación de las políticas de TI de una empresa.

CDO: Del inglés “Chief Digital Officer”, es un cargo surgido recientemente a raíz del desarrollo del comercio en línea. Es aquella persona que guía la transformación de ser una empresa física en ser una empresa digital, realiza las principales implementaciones de cambio en el medio digital, deben ser personas adaptables al cambio y también dirigen las tareas de los Community Manager.

Community Manager: Es la persona responsable por la comunidad en línea, es un profesional en marketing digital que permite la gestión y el desarrollo de una marca o una empresa en el mercado digital, así como aumentar la comunidad para detectar potenciales clientes y suscriptores.

1.6 Conclusiones

La inteligencia de negocios, realiza una valiosa herramienta para la organización ya que permite conocer el estado de sus procesos, pero más importante saber cuáles son los puntos críticos que debe atacar para garantizar el rendimiento ideal de su sistema.

Se logra evidenciar el aporte que tiene el esquema dimensional al facilitar el descubrimiento de los registros de eventos, y que posteriormente son analizados desde el modelo respecto al modelo de procesos organizacional, y permite identificar el estado de los procesos en función del comportamiento de las actividades, y el detalle del comportamiento de estas actividades que se encuentran en las tareas y variables provenientes de la base de datos transaccional.

De esta manera permite obtener una solución dinámica y flexible para la gestión de procesos de negocio, que mejora la eficiencia en las etapas de modelado, automatización, integración, monitoreo y optimización en forma frecuente.

En cuanto, la administración de empresas es una ciencia cuya teoría formulada data del siglo pasado; sin embargo, gracias a los avances tecnológicos ahora ha podido ser en ciertos campos reformulada ya que la inteligencia de negocios es un conjunto de ciencias interdisciplinarias donde las ciencias administrativas, las ciencias de la ingeniería de software y la novel ciencia de los datos muestran una importante sinergia.

Una de las principales falencias de la administración de las empresas era saber que la toma de decisiones era la correcta y que se podía aprovechar al máximo esta decisión; en cambio, en la actualidad con la herramienta de la inteligencia de negocios se dispone de una serie de métodos y técnicas, basadas en la información en función de la generación de conocimientos a través de sus componentes principales como la multi-dimensionalidad, minado de datos, agentes y almacén de datos. Todo ello les permite a los usuarios tener a disposición la información correcta en el lugar correcto, logrando una ventaja competitiva con respecto a su

competencia para contar con un conjunto de valores que contribuyan a la toma de decisiones, tanto desde el punto de vista estratégico como táctico.

Cuando comenzó la revolución del “Big Data” a finales de la década del noventa, la información paso a convertirse de un activo intrascendente a un activo muy valioso en la organización, sin embargo, no se debería sobrevalorar puesto que el objetivo de su significado no se refiere a la cantidad de datos, sino más bien a su importancia en cuanto a su transformación; por cuanto una vez transformado en conocimiento es sinónimo de revolución y evolución en las empresas.

1.7 Bibliografía

Alavi, M. & Leidner, D. (1999). *Knowledge management systems: issues, challenges communications of the association for Information Systems*. Obtenido de <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=2486&context=cais>

Ali, S. (2017). *Mastering ZOHO CRM, Manage your Team, pipeline, and clients effectively*. Apress Springer Nature

Anthony, P. (1990). The paradox of the management of culture or the who leads is lost. *Personnel Review*.

Ayas, K. (2001). *Estruturação de projetos para a aprendizagem e a inovação..* Sao Paulo: Editorial Atlas.

Baena, E. (11 de noviembre de 2009). *El Sector Público*. Recuperado de Aprendiendo economía: <https://aprendeconomia.com/2009/11/11/3-el-sector-publico/>

Bonales, J., y Zamora, A. (2015). Variables e índices de Competitividad de las empresas exportadoras, utilizando el PLS. *CIMEXUS* Vol. X, No. 2, 20.

Bonillo, P. (2004). *Evaluation system model for remote control software*. Serbiluz, 123-134.

- Bonnefoy, J., y Armijo, M. (2005). *Indicadores de desempeño en el sector público*. Santiago de Chile: Cepal.
- Brooking, A. (1997). *El capital intelectual. El principal activo de las empresas del tercer*. Barcelona: Paidós.
- Cañadas, E., y González, M. J. (2008). Los Indicadores de Gestión y el Cuadro de Mando en las Entidades no Lucrativas. CIRIEC - España, *Revista de economía pública, social y cooperativa*, 227-252.
- Chávez, V. P., y Pilco, J. P. (14 de junio de 2011). *Auditoría de Gestión al Departamento de Contabilidad del gobierno Municipal de Penipe correspondiente al período enero – diciembre 2010 a fin de alcanzar la eficiencia operativa*. Repositorio de la Escuela Superior Politécnica de Chimborazo - Dspace. Recuperado de <http://dspace.esPOCH.edu.ec/bitstream/123456789/2660/1/82T00135.pdf>
- Cleveland, W. (2001). Data science: an action plan for expanding the technical areas of the field of statistics. *International Statistical Review / Revue Internationale de Statistique*, 21-26.
- Conesa, J., y Curto, J. (2010). *Introducción al Business Intelligence*. Barcelona: Editorial UOC.
- Consejo de Judicatura. (2013). *Código Orgánico de la Función Judicial*. Quito: Registro Oficial .
- Cubero, J., y Berzal , F. (30 de 08 de 2011). *Sistemas Inteligentes de Gestión*. Recuperado de *Sistemas Inteligentes de Gestión* : <http://elvex.ugr.es/decsai/intelligent/workbook/ai/PROLOG.pdf>
- Davenport, T., & Prusak, L. (2001). *Working knowledge: How organizations manage what they know*. Boston: Harvard Business School Press. Boston: Harvard Business School Press. Buenos Aires: Pearson Education.

- Díaz, J. (15 de enero de 2017). *¿Cuál es el rol del administrador de empresas?*
Recuperado de Emprendices: <https://www.emprendices.com>
- Díaz, Y.; Pérez, Y., y Proenza, D. (2014). Sistema para la Gestión de la Información de Seguridad Informática en la Universidad de Ciencias Médicas de Holguín/
System for the Management of the Information of Informatics Security at the Medical Sciences University of Holguín. Holguín Ciencias, 1-14.
- Dixon, N. (2001). *Aprendendo através das fronteiras organizacionais*. Aprendizagem organizacional e organização de aprendizagem. Sao Paulo: Editorial Atlas.
- Dresner, H. (2010). *Profiles in performance: Business Intelligence Journeys and the roadmap for change*. United States: John Wiley & Son.
- Drucker, P. (2003). *Llega una nueva organización a la empresa. Gestión del conocimiento*. Bilbao: Ediciones Deusto.
- Burns, E. (20 de 07 de 2018). *TechTarget*. Consultado el 20 de 07 de 2018, de Infografía: La evolución del rol del director de datos. Recuperado de <https://searchdatacenter.techtarget.com/es/cronica/Infografia-La-evolucion-del-rol-del-director-de-datos>
- Escuela Europea de Management. (07 de marzo de 2017). *Capacitación y desarrollo personal en una empresa*. Recuperado de Desarrollo personal: <http://www.escuelamanagement.eu/>
- Espinoza, R. (08 de septiembre de 2016). *Indicadores de gestión - ¿Qué es un KPI?*
Recuperado de Welcome to the new Marketing: <http://robertoespinosa.es/2016/09/08/indicadores-de-gestion-que-es-kpi/>
- Feher & Feher. (2015). *Estandarizar operaciones beneficia el desarrollo de las empresas*. *El Financiero*, 61. Recuperado de [HYPERLINK "https://www.elfinanciero.com.mx/empresas/estandarizar-operaciones-beneficia-al-desarrollo-de-empresas-feher-feher"](https://www.elfinanciero.com.mx/empresas/estandarizar-operaciones-beneficia-al-desarrollo-de-empresas-feher-feher)

<https://www.elfinanciero.com.mx/empresas/estandarizar-operaciones-beneficia-al-desarrollo-de-empresas-feher-feher>

Fernández, A., y Llorens, F. (2017). Gobierno de las TI para Universidades. *Repositorio Institucional de la Universidad de Alicante*.

Flores Jimeno, M. D., y Navarro Galera, A. (2010). *Los Indicadores de Gestión de las Administraciones Públicas*. Andalucía: Universidad de Granada.

Gartner. (16 de 02 de 2017). *Magic quadrant business intelligence analytics*. Recuperado de <https://www.gartner.com/doc/3611117/magic-quadrant-business-intelligence-analytics>

Gartner, N. (13 de 12 de 2017). *Gartner*. Obtenido de Gartner Says By 2020, Artificial Intelligence Will Create More Jobs Than It Eliminates. Recuperado de <https://www.gartner.com/newsroom/id/3837763>

Gelard, P. B. (2014). Relationship between Transformational Leadership and Knowledge Management. *International Journal of Information Science and Management*, 12 (2), 67-82. .

George, P. (2015). *Teach Yourself Visually WordPress*. 3rd Edition.

GNU (s.f.). *General Public License*. Recuperado de [HYPERLINK "https://www.gnu.org/licenses/licenses.es.html"](https://www.gnu.org/licenses/licenses.es.html)
<https://www.gnu.org/licenses/licenses.es.html>

Guinart, J. M. (2003). Indicadores de gestión para las entidades públicas. VIII *Congreso Internacional del CLAD sobre la Reforma del Estado y DE LA Administración Pública*, Panamá, 28-31.

Guinart i Solá, J. M. (2003). *Indicadores de Gestión para las Entidades Públicas*. Panamá: Escuela de Administración Pública de Catalunya.

- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques (Vol. 3th Edition)*. Massachusetts: Elsevier.
- Hatch, M. (1997). *Organization theory*. Oxford:University Press.
- Hernández, R. (20 de junio de 2010). *Establecer Ilos Indicadores de Gestión como herramienta para optimizar los procesos en cada uno de los subsistemas de la oficina de Recursos Humanos del Ministerio de Poder Popular para la Energía y Petróleo*. Universidad Simón Bolívar - Vicerrectorado Académico. Recuperado de <http://159.90.80.55/tesis/000148811.pdf>
- Holthan, C. (1998). *Culture, not IT, the barrier to success of KM initiatives, says IBM*. Knowledge Management.
- Holthan, C. (1998). *Keeping a weather eye on the big picture*. Knowledge Management, pp. 21-23.
- Indurkha, N., & Damerau, F. (2010). *Handbook of Natural Language Processing*. Boca Raton Florida: CRC Press.
- Integra . (16 de diciembre de 2015). *Consultores de Sistema de Gestión*. Recuperado de <http://www.consultoresdesistemasdegestion.es/sistemas-de-gestion/>
- Jeff, B. (2017). *Business in Real-Time Using Azure IoT and Cortana Intelligence Suite*.
- Kaplan, J. (2017). *Inteligencia Artificial: Lo que todo el mundo debe saber(R)*. Madrid: Teell Editorial.
- Kimball, R. (2004). *The Data Warehouse ETL Toolkit*. Wiley.
- Kragh, P. (2009). *Picture this: Managed change and resistance in business network*.

- Lastra Calderón, N. P. (04 de mayo de 2011). *Diseño de Indicadores de Gestión para la empresa pública municipal de Rastro de Tulcán, basado en el modelo del cuadro de mando integral*. Repositorio de la Universidad Técnica Particular de Loja - Dspace. Recuperado de <http://dspace.utpl.edu.ec/bitstream/123456789/4784/1/TesisNixonLastra.pdf>
- Laudon, J.P.y Laudon, K.C. (2012). *Sistemas de información gerencial* (Vol. 12). EE.UU: Pearson.
- Laudon, J.P. (2004). *Sistemas de Información Gerencial*. México: Prentice Hall.
- López , A., & Ortiz, D. (2014). Los Indicadores de gestión y el control de la eficiencia del sector público. *Revista Española de Control Externo*, 189 - 218.
- Lopez, M., y Guerrero , D. (2017). *Modelo de Inteligencia de Negocios y Analítica en la nube para PYMES del sector retail*. Recuperado de: <http://revistas.utp.ac.pa/index.php/memoutp/article/view/1466>
- Luhn, P. (1958). A Bussiness Intelligence System. *IBM Journal*, p. 314.
- Mamaghani, F. (2002). *Information Technology Knowledge Sharing Using case-Based*. Informarion Systems Management.
- Martínez, J. (2012). *El New Public Management*. París: Institutno Internaacional de Ciencias Administrativas .
- Mayo, E. (1972). *Problemas humanos de una civilización industrial*. Buenos Aires: Nueva visión.
- Mayo, E. (1972). *Problemas humanos de una civilización industrial*. Buenos Aires: Nueva visión.
- Microsoft. (05 de 05 de 2018). *Conceptos de minería de datos*. Recuperado de: <https://docs.microsoft.com/es-es/sql/analysis-services/data-mining/data-mining-concepts?view=sql-analysis-services-2017>

- Moral, D. (2007). *Gestión del conocimiento*. España: Thomson Editores.
- Nihad A Hassan, R. H. (2018). Open source Intelligence Methods and tools a practical guide to on line intelligence.
- Nonaka, I., & Takeuchi, H. (1999). *La organización creadora de conocimiento*. México: Oxford University Press.
- OECD. (2003). Measuring Knowledge Management in the Business Sector. First Steps [en línea].
- Ogbonna E. & Wilkinson, B. (1990). Corporate strategy and corporate culture: the management of change in the UK supermarket industry. *Personel review*.
- Oquendo, M. del P. (2011). *Indicadores de Gestión*. Popayán: Alcaldía de Popayán.
- Ordoñez de Pablos, P. (2003). Marco conceptual para el análisis del conocimiento. Alta Dirección.
- Orozco, M., y Quiroz, G. (19 de Julio de 2015). Diario el Comercio. El 82% de pymes de Ecuador accede a Internet, pero su uso se limita a enviar correos y tareas administrativas, p. 1.
- Ouchi, W. (1982). Teoría Z: como pueden las empresas hacer frente al desafío japonés. Bogota: Norma.
- Paredes, J. (21 de junio de 2011). *Diseño de un Sistema de Indicadores de Gestión en la División de Auditoría Técnica de FONTUR para el seguimiento y control de ejecución de obras públicas*. Recuperado de Repositorio de la Universidad Católica Andres Bello:
<http://biblioteca2.ucab.edu.ve/anexos/biblioteca/marc/texto/AAS2279.pdf>
- Peters, T. (1984). *En búsqueda de la excelencia*. Bogotá: Norma.

- Poggio, J. (2013). It management model for financial report. *Journal of Information Systems and Technology Management*, 597-620.
- Rodríguez, M. (2009). La gestión del conocimiento, factor estratégico para el desarrollo. *Revista de Investigación de la Facultad de Ciencias Administrativas*.
- Ruíz, L. M. (2016). Roles y estilos gerenciales en las cooperativas multiactivas de la ciudad de manizales. *Summa Iuris*, 3(2), 405-440.
- Saur, B. (1998). *Library and Information Science Abstracts*. Recuperado de
HYPERLINK
"https://biblioteca.ugr.es/pages/biblioteca_electronica/bases_datos/science-abstracts-lisa"
https://biblioteca.ugr.es/pages/biblioteca_electronica/bases_datos/science-abstracts-lisa
- Siegel Clifford, F. D. (2014). A Guide to Delivering Business Results with Big Data Fast Actionable Intelligence.
- Siegel, E. (2013). *Análisis predictiva: Predecir el futuro utilizando Big Data*. Madrid: Anaya Multimedia.
- Solarte, F., Enríquez, E., y Benavides, M. (2015). Metodología de análisis y evaluación de riesgos aplicados a la seguridad informática y de información bajo la norma ISO/IEC 27001. *Revista Tecnológica ESPOL*, 492-507.
- Steven, F. (2018). The Modern Approach to Continuous Performance Improvement for Businesses Optimizing Data-to-Learning-to-Action.
- Suresh, J.K & Mahesh, K (2006). Ten Steps to Maturity in Knowledge Management. Chandos Publishers
- Weiss, S., & Indurkha, N. (1997). Predictive Data Mining. Morgan Kaufmann.

Wrike.solution. (2017). *Silicon Valey*. Obtenido de

https://www.wrike.com/es/?ga_campaign=LATAM+SRCH+Brand+ES&ga_adgroup=Wrike&ga_keyword=wrike&targetID=aud-176471549087:kwd-12217897514&gclid=Cj0KCQjw3qzzBRDnARIsAECmryrEVVS3HffiULco4ILZUOwUBD8ud6bd9OXilzezsie_p7hrbhqMam0aAutPEALw_wcB

Yaghoubi, H. M. (2014). Transformational Leadership: Enabling Factor of Knowledge Management Practices. *Journal of Management and Sustainability*, 4 (3), 165-174.

Zimmer, M. (2005). A criação de conhecimento nas empresas. En Os novos horizontes da gestao:Aprendizagem organizacional e competencias. Brazil: Bookman.

1.8 ACTIVIDADES DE REPASO

- 1) La Inteligencia de negocios es el conjunto de:
 - a) Estrategias, aplicaciones, datos, productos, tecnologías que están enfocados a la creación y administración del conocimiento, a través del análisis de los datos.
 - b) Paradigmas que permiten generar conocimiento a gran escala.
 - c) Búsqueda de información sobre las bases de datos que permiten producir conocimiento propio a las empresas.
 - d) Conjunto de relaciones entre las diferentes áreas de una organización que permiten generar conocimiento específico en la organización.

- 2) El proceso de minería de datos difiere del proceso de almacén de datos:
 - a) Verdadero
 - b) Falso

- 3) La minería de datos es el proceso de:
 - a) Descubrir conocimiento mediante el uso de bases de datos corporativas.
 - b) Descubrir patrones en grandes volúmenes de datos almacenados en múltiples almacenes de datos.
 - c) Encubrir datos en grandes volúmenes de datos para posteriormente realizar una encriptación de los mismos.
 - d) Esconder datos para después usarlos en base de datos corporativas.

- 4) ¿Cuáles son las principales aplicaciones del análisis predictivo?

- 5) El director de datos desempeña el cargo de:
 - a) Responsable del área de base de datos, resaltando su habilidad para manejar la programación en la base.

- b) Responsable del manejo de las comunicaciones de datos, entre e inter organización, comunicando las novedades al Gerente General.
- c) Responsable de tomar decisiones basadas en hechos y evidencias.
- d) Responsable de manejar las tecnologías de la información de la organización.

6) Mencione tres formas de aplicar la minería de texto.

- a) _____
- b) _____
- c) _____

7) El lenguaje de procesamiento natural es un campo:

- a) Del lenguaje de programación que estudia el procesamiento natural entre el ambiente y las máquinas.
- b) De computación que estudia el lenguaje para procesar (en forma natural) la lingüística entre las personas y las máquinas.
- c) De los lenguajes de programación que verifica la capacidad de procesar (en forma natural) las personas y la aplicación de control.
- d) De las ciencias de la computación y la lingüística que estudia las relaciones e interacciones entre las computadoras y el lenguaje natural.

8) Los tableros de control son:

- a) Indicadores que permiten verificar el estado actual de las computadoras y sus aplicaciones en forma de métricas.
- b) Herramienta de inteligencia de negocios que permite mostrar el estado actual de las métricas y los indicadores claves de desempeño.
- c) Herramienta de la minería de datos que permite visualizar el estado actual de los almacenes de datos y sus indicadores claves de desempeño.
- d) Indicadores y métricas que permiten verificar el estado actual de los indicadores de la organización en sus diferentes áreas.

9) ¿Cuál es la diferencia entre OLAP y ETL?

10) Mencione tres razones para crear un almacén de datos.

- a) _____

- b) _____

- c) _____

11) Un KPI o indicadores claves de desempeño es una medida del nivel de:

- a) Desempeño de un dato en particular.
- b) Desempeño de una aplicación clave de desempeño.
- c) Desempeño de un grupo humano y sus aplicaciones clave.
- d) Rendimiento de un proceso, que está relacionado previamente con un objetivo.

12) Las herramientas estadísticas en la minería de datos son:

- a) Histogramas, Herramientas de tendencia (media, varianza) y Distribución binomial.
- b) Distribuciones estadísticas en general.
- c) Algoritmos de asociación, de clustering y de clasificación.
- d) Hipótesis de una y dos colas.

13) La utilización del KPI para crear un DASHBOARD es la representación gráfica de:

- a) Los principales KPI para la consecución de los objetivos en la organización.
- b) Los objetivos específicos para conseguir los objetivos generales en la organización.
- c) Los números relevantes de cada área funcional de una organización.
- d) Los principales activos y pasivos que tiene la organización para conseguir los objetivos principales.

14) Las secciones principales de SCORECARD son:

- a) Mapa conceptual, medidores de calidad, las metas y las iniciativas.
- b) Mapa estratégico, indicadores de gestión, las metas, las iniciativas, proyectos estratégicos.
- c) Mapa conceptual, indicadores de gestión, los objetivos y las iniciativas.
- d) Mapa estratégico, medidores de calidad, los objetivos y las iniciativas.

15) ¿Qué es un analista de datos?
