

Análisis bibliométrico de la producción científica sobre la aplicación de las ciencias de datos para la toma de decisiones en análisis de riesgos, 2015-2020

Bibliometric análisis of scientific production on the application of data science for decisión making in risk análisis, 2015-2020

ARANGO, Juan F.¹
DÍAZ, Joan S.²
SANCHEZ, Juan M.³

Resumen

Este artículo presenta una revisión a los trabajos desarrollados en la aplicación de métodos de ciencia de datos para calcular y gestionar los riesgos financieros de los proyectos de ingeniería, enfocándose en la simulación y los resultados obtenidos, analizando los datos históricos, países, autores, temática y sectores en los cuales se ha desarrollado el objeto de estudio, mediante las herramientas especializadas VOS viewer y Scopus, encontrando cómo es un tema nuevo que viene tomando relevancia en los últimos años.

Palabras clave: ciencias de datos, análisis bibliométrico, análisis de riesgos, gestión empresarial

Abstract

This article presents a review of the work carried out in the application of data science methods to calculate and manage the financial risks of projects, focusing on simulation and the results obtained, analysing the history, countries, authors, subject matter and sectors in which it has been developed in the object of study, finding the main difficulties in the application of data science in financial risk management.

key words: data science, bibliometric analysis, risks analysis, business management

1. Introducción

La toma de decisiones es un proceso fundamental en el desarrollo de la gestión de un proyecto, por ello se ha visto un incremento en los distintos métodos dentro de los análisis cuantitativos y la administración de empresas (Amaya, 2010); también existen diferentes perspectivas que analizan esa toma de decisiones y su impacto, introduciendo conceptos como el círculo de las decisiones (Brenes, 2016) que analiza no solo la decisión a tomar sino la velocidad con la que se toma esa decisión y la diferencia entre una decisión “importante” y una de

¹ Estudiante. Ingeniería de sistemas. Universidad Distrital Francisco José de Caldas. jfarangom@correo.udistrital.edu.co

² Estudiante. Ingeniería de sistemas. Universidad Distrital Francisco José de Caldas. josdiazp@correo.udistrital.edu.co

³ Ingeniero Electrónico, Especialista en Teleinformática. Magister en Administración. Profesor Asociado e integrante del grupo de investigación GIIRA y AQUAFORMAT. Universidad Distrital Francisco José de Caldas. Facultad de Ingeniería. jmsanchezc@udistrital.edu.co

“rutina”. Este panorama permite ver el modo en que la toma de decisiones es un proceso crítico en el que no hay nada escrito; para ello en los últimos años gracias a la ciencia de datos, al creciente número de información dentro de las bases de datos de las organizaciones y a los cambios tecnológicos ha dado origen a la necesidad u oportunidad de analizar esos datos (Aguilar, 2016) y con la información que surge a partir de ese análisis comenzar el proceso de la toma de decisiones. Un ejemplo de esto es la creciente ola de técnicas de minería de datos, y cómo estas tienen cada vez más cabida en los campos de la ciencia, y “cabe destacar, en las aplicaciones financieras y en banca, en análisis de mercados y comercio, en seguros[...]” (Pérez y Santín, 2008)

Con esto, se logra tener una noción de las tendencias en los métodos para la toma de decisiones y sobre todo cuál es el papel fundamental que juegan las bases de datos en la generación de esa información que apoya el proceso de las decisiones; hoy en día existen diversas maneras de analizar el riesgo financiero; desde métodos como el análisis de estados financieros o el análisis de índices financieros, hasta métodos no tradicionales como el método Delphy (Maozhu Jin, 2018); algunos autores indican que “quienes deben decidir entre las diversas opciones de inversión o quienes deben sugerir la movilización de recursos hacia un determinado proyecto, asumen una gran responsabilidad, pues sus recomendaciones pueden afectar en forma significativa los intereses de los inversionistas” (Miranda, 2005) lo que conlleva que se prefiera una tener la mayor cantidad posible de información de apoyo para asegurar que la decisión tomada sea la ideónea.

Dadas las recientes necesidades en el crecimiento de gestión de riesgos de una empresa, se han buscado alternativas a la hora de evaluarlos, tal que dicha evaluación sea eficiente, exhaustiva y objetiva (Maozhu Jin, 2018); es por esto que en este artículo se tiene como objetivo realizar una revisión de los trabajos que se han desarrollado dentro del campo de la gestión de proyectos donde se enfatice la toma de decisiones con grandes cantidades de datos, junto a las herramientas utilizadas y los resultados que se han obtenido.

Para ello, el artículo se organiza de la siguiente manera: La primera sección busca presentar el marco teórico que permite conocer a grandes rasgos las definiciones necesarias para comprender qué es la ciencia de datos, el análisis de riesgos dentro de la gestión de proyectos y los riesgos financieros, luego se presenta la metodología usada y los principales hallazgos de la investigación; y finalmente, se encuentran los resultados obtenidos, su respectivo análisis y las conclusiones resultantes de lo anterior.

2. Marco teórico

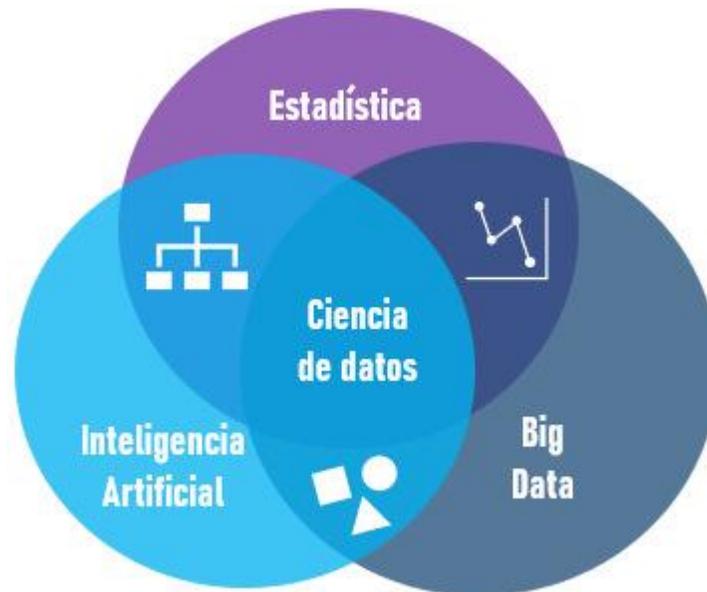
Los conceptos desarrollados a continuación son necesarios para la comprensión del tema abordado en este artículo. Primero se encuentra la definición de la ciencia de datos con los principales campos que la componen como la Estadística, la inteligencia artificial y el análisis de datos o Big data; posteriormente se define el análisis de riesgos dentro de la gestión de proyectos.

2.1. Ciencia de datos

La ciencia de los datos es un área de trabajo interdisciplinar en la que se puede distinguir procesos como la recopilación, preparación, análisis, modelado y presentación de datos que resulten en conocimiento útil para la interpretación de problemas que no son fácilmente comprensibles, ayudando así en la toma de decisiones (Provost y Fawcett, 2013).

Para ello, la ciencia de datos ha recurrido a distintas áreas del conocimiento que le permiten alcanzar sus objetivos.

Figura 1
Representación de las áreas del conocimiento
empleadas por la ciencia de datos.



Fuente: Elaboración propia, basada en <https://intellipaata.com/blog/what-is-data-science/> (IntelliPaata, s.f.)

Como se muestra en la figura 1, la ciencia de los datos es la combinación de 3 disciplinas fundamentales que llevadas a cabo de manera correcta permiten realizar un análisis completo sobre los datos para generar información que apoye los procesos de toma de decisiones.

En los últimos años, esta disciplina se ha estado fortaleciendo y desarrollando, en parte gracias a tres factores fundamentales: la gran cantidad de datos que se genera día a día y la disponibilidad de los mismos, las sofisticadas tecnologías que crecen a la par que la industria y la sociedad misma, y finalmente la potente infraestructura computacional con la que se cuenta; esto genera la oportunidad de obtener y generar la información necesaria para mejorar las áreas de la sociedad, desde la educación y los negocios hasta la política (Agarwal y Dhar, 2014).

El resultado de la ciencia de datos no está propiamente definido, es más bien la información resultante de un análisis del problema en cuestión, utilizando la combinación de las disciplinas mencionadas anteriormente; por esto los resultados de una análisis pueden indicar que features o categorías tienen más influencia en cada contexto, pero se puede obtener un resultado contrario si se implementa un clasificador con redes neuronales; pues se tendría una baja interpretación de cómo funciona el modelo y esto puede llevar a que se desconozca de manera verídica qué feature aporta más información, además se podría implementar una metodología cíclica en la que en cada iteración se busque una mayor exactitud o una en la cual sea necesario encontrar un mejor desempeño en la generación de predicciones. Cada técnica y cada metodología generará un resultado diferente, dependerá del equipo de investigación o del desarrollo del enfoque que desean y el valor que se le dará a cada resultado por lo que la gama de resultados crece en función a los nuevos desafíos e incógnitas de la sociedad actual. (Aalst, 2016).

2.1.1. Estadística

La ciencia de datos se compone de estadística y matemáticas, en donde se observa que el estado actual del trabajo relacionado a la estadística puede resumirse en tres pilares fundamentales: Primero, recolección de datos;

segundo, modelado y análisis; y tercero, la comprensión del problema y la toma de decisiones para su posterior solución (Wu).

Las herramientas de regresión y clasificación como los métodos de Bayes, Laplace o Gauss y posteriormente en el siglo 20, el desarrollo de los conceptos de ideas de diseño experimental y estimación de máxima probabilidad dieron vida a las regresiones, concepto clave dentro de la ciencia de los datos, pues estas permiten generar funciones matemáticas por medio de técnicas de aprendizaje de máquina, que a partir de un conjunto de datos conocidos logra predecir el comportamiento o clasificarlo (Bruce y Gedeck, 2017).

Los estadísticos desarrollan ahora trabajo de análisis de datos, y algunos autores (Cleveland, 2001) sugieren que este es el nuevo camino al cual debe apuntar la estadística; pues dentro de la ciencia de los datos el proceso de limpieza o transformación de los datos requiere manejos matemáticos y particularmente, los estadísticos.

2.1.2. Inteligencia artificial

Se define generalmente como el desarrollo de software o máquinas que presentan un comportamiento comparable con el de los seres humanos (Russel Y Norving, 2008), para tener un marco de referencia. Se han concretado cuatro enfoques desde los cuales se han abordado las áreas subyacentes de este campo; dichos enfoques son:

Figura 2
Enfoques de la inteligencia artificial

Sistemas que piensan como humanos	Sistemas que piensan racionalmente
<p>«El nuevo y excitante esfuerzo de hacer que los computadores piensen... máquinas con mentes, en el más amplio sentido literal». (Haugeland, 1985)</p> <p>«[La automatización de] actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas, aprendizaje...» (Bellman, 1978)</p>	<p>«El estudio de las facultades mentales mediante el uso de modelos computacionales». (Charniak y McDermott, 1985)</p> <p>«El estudio de los cálculos que hacen posible percibir, razonar y actuar». (Winston, 1992)</p>
Sistemas que actúan como humanos	Sistemas que actúan racionalmente
<p>«El arte de desarrollar máquinas con capacidad para realizar funciones que cuando son realizadas por personas requieren de inteligencia». (Kurzweil, 1990)</p> <p>«El estudio de cómo lograr que los computadores realicen tareas que, por el momento, los humanos hacen mejor». (Rich y Knight, 1991)</p>	<p>«La Inteligencia Computacional es el estudio del diseño de agentes inteligentes». (Poole <i>et al.</i>, 1998)</p> <p>«IA... está relacionada con conductas inteligentes en artefactos». (Nilsson, 1998)</p>

Fuente: Russell. Stuart. NORVIG. Peter. Inteligencia Artificial, Un enfoque moderno. 2008. (Russel y Norving, 2008)

Dentro de la IA (inteligencia artificial) tenemos dos líneas de trabajo principales dentro del contexto abordado:

Sistemas basados en conocimiento: son programas informáticos que razonan a partir del conocimiento previo que puede ser dado por un experto (Brock, 2018); dentro de este podemos encontrar los sistemas basados en reglas, los cuales nos permiten establecer condiciones de relación entre eventos determinados usando reglas del tipo If-Then o If-Then-Else; y los sistemas de razonamiento basados en casos, en donde los sistemas se encargan de tomar decisiones a partir de la experiencia de casos anteriores (Russel Y Norving, 2008).

Aprendizaje de máquina o Machine Learning: es una rama de los algoritmos de la inteligencia artificial orientada a soportar decisiones en contextos críticos en donde es importante la imparcialidad, enfocados al tratamiento de grandes volúmenes de datos (Barredo et al., 2018). “Las técnicas basadas en el aprendizaje automático se han aplicado con éxito en diversos campos que van desde el reconocimiento de patrones, la visión por

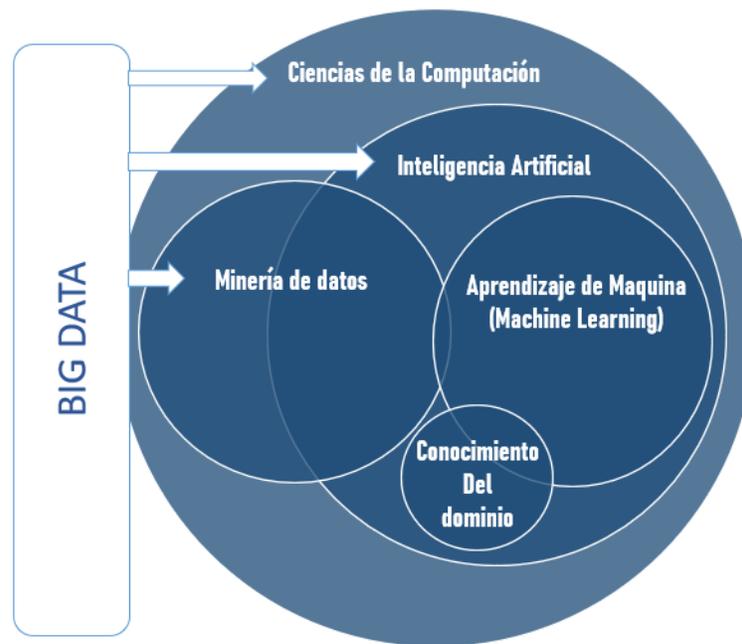
ordenador, la ingeniería de naves espaciales, las finanzas, el entretenimiento y la biología computacional hasta las aplicaciones biomédicas y médicas” (Lledó, 2007).

2.1.3. Análisis de datos y big data

“El concepto de Big data se puede definir como las múltiples fuentes de información de alto volumen, alta velocidad y alta variedad que exigen de formas innovadoras y costo efectivas para ser procesadas con el fin de generar descubrimientos, procesos de decisión y automatización de procesos” (Gartner, 2020).

Como se muestra en la figura 3, el análisis de datos se encuentra por encima de las dos áreas mencionadas anteriormente siendo esta, el punto más alto de las ciencias de datos.

Figura 3
Enfoques de la inteligencia artificial

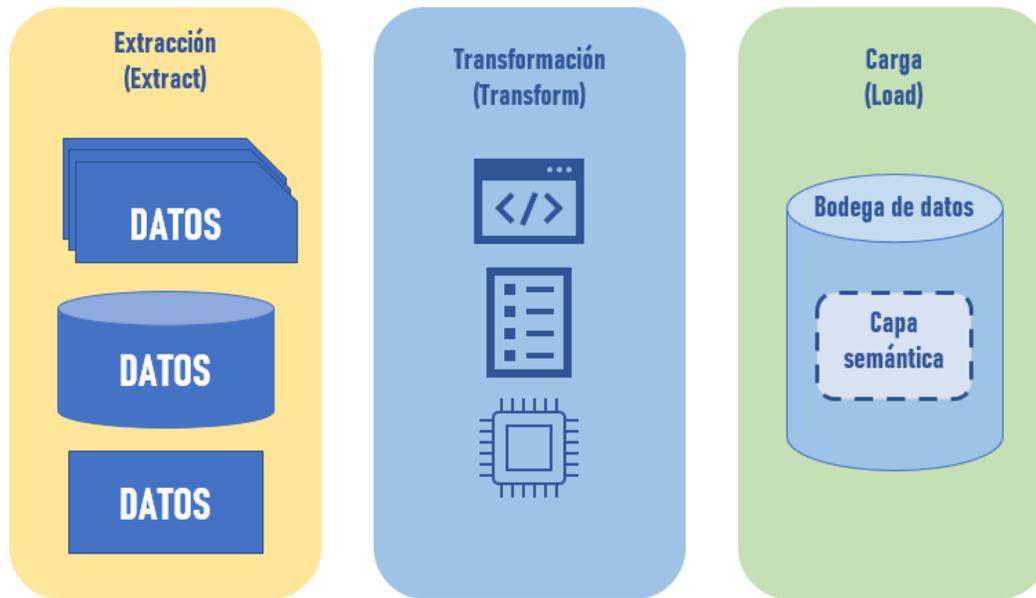


Fuente: elaboración propia, basada en Mora. J.
¿Qué es la ciencia de datos y cuál es su uso? (Mora, s.f.)

En la figura 3, vemos cómo la Big Data es un área del conocimiento que atraviesa varias disciplinas englobadas en el contexto de las ciencias de la computación, ya que la capacidad de procesamiento de las computadoras actual es la que permite transformar y analizar grandes cantidades de datos; y ya dentro de estas, la inteligencia artificial es la que más ha avanzado en el tema.

Con estos tres componentes ya definidos, se puede establecer el proceso fundamental de la ciencia de los datos, abreviado como ETL.

Figura 4
Descripción del proceso ETL



Fuente: Elaboración propia basada en <https://www.jet-software.com/en/data-integration-fast-and-easy-iri-etl-tools/etl-diagram/> (Russell Kelly, s.f.)

En el proceso descrito en la figura 4, podemos ver la manera en que la Big data forma parte de la primera etapa; pues dentro de sus características, encontramos lo comúnmente denominado “Las 5 V’s”: volumen, velocidad, variedad, veracidad y valor (Lengauer, 2020); dichas características buscan la extracción de datos que son válidos para nuestro problema o contexto, esto se logra usando tecnologías como bodegas de datos, que permiten reunir información proveniente de diversos orígenes, información estructurada ya existente en bases de datos por lo general modelos relacionales, semi estructurada como archivos XML e incluso información no estructurada como archivos de texto o en formato PDF, para obtener la mayor y mejor cantidad de datos posible. Luego de esto se pasa al proceso de transformación que consiste en limpiar y convertir los datos para estandarizarlos, eliminar datos erróneos o vacíos, datos no coherentes, esto con el fin de tratarlos de la misma manera, para esto se usan sistemas de software especializados, o algunas librerías de procesamiento y preprocesamiento de datos, finalmente los datos se cargan en la bodega de datos o data Waterhouse para su subsiguiente análisis (Mukherjee y Kar, 2017).

Posteriormente, el proceso de ciencia de datos se basa en utilizar herramientas de ciencia de datos; y en el conocimiento del dominio específico en el que se está trabajando, para analizar la información. Para esto existen diversas tecnologías y metodologías, que generan datos útiles a los procesos de la empresa y que sean de ayuda en la toma de decisiones. (Menichini, 2004).

2.2. Análisis de riesgos

Dentro de la formulación y la gestión de proyectos de ingeniería el análisis de riesgos es uno de los estudios que se requiere para decidir si es factible o no, mucho más dentro de los proyectos de inversión que tienen “como propósito demostrar que los proyectos examinados no solo son factibles de ejecutar desde el punto de vista técnico, sino que, además, pueden pagar el costo de la inversión y generar una ganancia adicional” (Espinoza, 2010). Adicional a esto, el PMBOK (estándar norteamericano para la gestión de proyectos) lo considera como una de las áreas de la gestión de proyectos, la gestión y el análisis de los riesgos (Assaff, 2006) mencionando que este proceso es vital para el éxito o fracaso de los proyectos.

2.2.1. Riesgo financiero

Se define el riesgo financiero como aquel riesgo que proviene de las pérdidas dentro de los balances provenientes del movimiento de los precios del mercado, y se relaciona con la incertidumbre de las inversiones generadas y las obligaciones financieras que el proyecto no pueda cumplir. Esto se puede definir en un único valor, los costos fijos en que incurre el proyecto (Mukherjee y Kar, 2017) (Padilla, 2012).

La variabilidad en los resultados distintos a los esperados puede tener múltiples causas y dependiendo de ellas, puede clasificarse en: riesgo de mercado, riesgo de crédito, riesgo de liquidez, riesgo operacional o riesgo legal (Menichini, 2004).

Las variables internas del proyecto hacen referencia a aquellas que se pueden controlar, y pueden verse afectadas por variables externas que pueden ser: ingresos, otros ingresos, gastos, costos, nivel de financiación, nivel de inversión, tasa de descuento apropiada por parte del inversionista, rotación de cuentas por cobrar, rotación de inventarios, rotación de cuentas por pagar, tasa de servicio o comercialización, distribución de utilidades y niveles mínimos de caja para operar.

Como variables externas tenemos a aquellas que no se pueden controlar y son de índice macroeconómico, en Colombia, las más predominantes son: las tasas de interés local: (DTF, TCC, TES); acciones y fondos fiduciarios, entre otros; las tasas de inflación, la tasa de cambio, PIB, niveles de ahorro, niveles de inversión, tasas de crecimiento sectoriales, tasas de impuestos, riesgos de inversión en el sector, riesgos de inversión en el país y el riesgo de inversión en la ciudad o localidad.

3. Herramientas de análisis de riesgos

Para desarrollar un análisis de riesgos existen diferentes técnicas y herramientas, tanto cualitativas como cuantitativas.

3.1. Análisis cualitativo de riesgos

Este tipo de análisis trata de comprender los riesgos y el impacto de estos a través del conocimiento que se tiene del contexto en el que se desarrolla el proyecto, como herramientas para dicho análisis se encuentra la tormenta de ideas (brainstorming) o entrevistas a los miembros del equipo o el método Delphi (Lledó, 2007).

3.1.1. Método delphi

Es un método de estructuración y modelo de comunicación para grupos que busca tratar a cada uno de los miembros del mismo como un todo y permitir tratar problemas complejos a partir de hacer estimaciones futuras sobre un grupo de expertos en el tema a tratar. “la capacidad de predicción de la Delphi se basa en la utilización sistemática de un juicio intuitivo emitido por un grupo de expertos.” (Hipodec, 2018)

3.2. Análisis cuantitativo de riesgos

Dentro de los métodos cuantitativos, se busca estimar la probabilidad de que un suceso ocurra, teniendo en cuenta su grado de impacto respecto a las variables de riesgo. Las principales técnicas para realizar el análisis cuantitativo exigen la recolección de datos, la aplicación de técnicas cuantitativas, y técnicas de modelamiento. Las técnicas de análisis cuantitativo más utilizadas son: el análisis de árboles de decisión y la simulación (Del Carpio, 2006).

Los árboles de decisión son un conjunto de diferentes algoritmos que permite representar la información en manera de árbol; son una de las técnicas más utilizadas por su practicidad y este se construyen a través del entrenamiento de datos clasificados; por lo que son algoritmos supervisados y generar un árbol donde los nodos intermedios son los atributos o características del objeto a clasificar y las hojas son las clasificaciones; esto se

consigue utilizando los algoritmos que mejor clasifiquen, es decir, que proporcionen más información (Britos, 2006).

3.2.1. Simulación

A partir de las variables del proyecto, tanto internas como externas, se pueden realizar simulaciones para generar diversos escenarios a través de distintos métodos para así poder prever los riesgos del proyecto y tener planes para gestionarlos. (Bazzani y Cruz, 2008).

3.2.1.1. Método montecarlo

La simulación de Monte Carlo es una técnica que utiliza conceptos de estadística y de computación para la generación de números pseudo-aleatorios y cálculos automatizados.

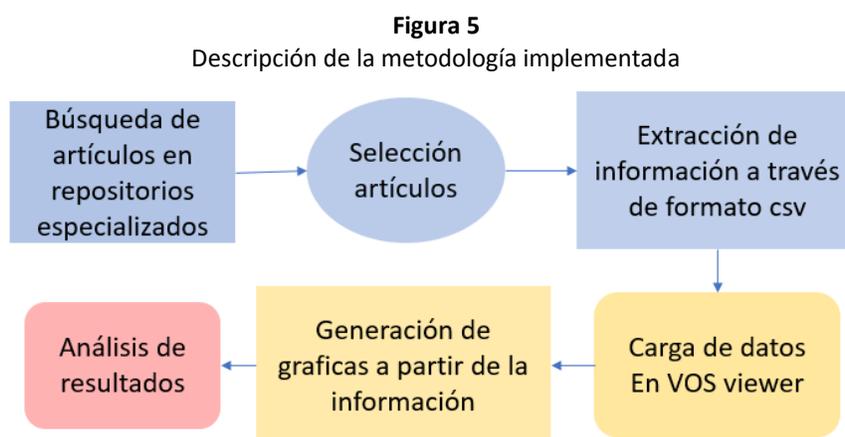
Se basa en el uso de modelos matemáticos que describen el comportamiento de un proceso o actividad, identificando las variables de entrada; una vez identificadas esas variables de entrada que afectan el comportamiento global del sistema se lleva a cabo simulaciones con los valores pseudoaleatorios generados, esto se repite n veces, lo que genera n observaciones del comportamiento del sistema, mientras mayor sea en n mayor será el entendimiento del sistema (Faulín y Juan, 2005).

3.2.1.2. Análisis de estados financieros

Consiste en reconstruir las actividades cooperativas y de todas las transacciones de contabilidad, análisis del estado de los flujos contables apoyado por sistemas informáticos para el almacenamiento de datos, acceso y revisión de grandes cantidades de datos, revisando los principios de contabilidad, enfoque y propósito (Bernstein, 1997).

4. Metodología

El desarrollo de este artículo se realizó usando dos herramientas para la obtención y análisis de información bibliográfica de la temática abordada. La primera es SCOPUS, la cual es un índice bibliográfico que contiene una colección representativa, completa y multidisciplinar a nivel mundial; de esta herramienta se recuperaron 331 artículos entre el año 2010 y 2019 que fueron resultado de la búsqueda "Artificial intelligence" and "Project management", usados como apoyo de la investigación. La segunda herramienta bibliométrica es VOSviewer desde la que se obtuvo las redes de coincidencia entre los tópicos que tienen mayor importancia dentro de las investigaciones relacionadas al tema principal, junto con los mapas de densidad de cada área, en donde se tomaron las palabras clave indexadas y las propuestas por el autor.



Fuente: Elaboración propia

5. Resultados

Los resultados mostrados a continuación hacen parte del análisis realizado a los artículos relacionados con el tema.

5.1. Documentos

En la figura 6 se muestra el número de documentos publicados por año, donde se aprecia que los estudios sobre el tema han venido incrementándose de manera exponencial en el último lustro y se espera que esta tendencia se mantenga, se puede ver que año tras año las tecnologías de la información y el análisis de datos han estado cada vez más presentes en la toma de decisiones y el análisis de riesgos financieros como se evidencia en los trabajos (Muhammed y Kuklakli, 2018) (Haque et al, 2018) que indican que las empresas e investigadores centrados en las soluciones de datos, pueden incursionar en la recopilación y análisis de datos que permitan abordar el ámbito de riesgos financieros.

Figura 6
Crecimiento de investigación en el campo



Fuente: elaboración propia basada en datos de Scopus

Como se aprecia en la figura 7, si se revisa el tipo de recurso, se ve una marcada tendencia hacia los artículos indexados y los artículos presentados en conferencia enfocados al uso de las nuevas tecnologías en distintos campos de la industria, desde cuidados médicos (Shabaz, 2019) hasta el análisis de microcréditos (Islam y O’Gorman, 2019). Todo esto enfocado a la toma de decisiones y gestión empresarial.

Figura 7
Tipos de documentos



Fuente: Elaboración propia basada en datos obtenidos de Scopus

Revisando los resultados agrupados por país de publicación se puede ver una alta densidad en Estados Unidos e Inglaterra como se aprecia en la figura 8, mostrando el potencial que tienen dichos países en la materia siendo estos los que se encarguen de mostrar las directrices que permitan la innovación y desarrollo de este campo.

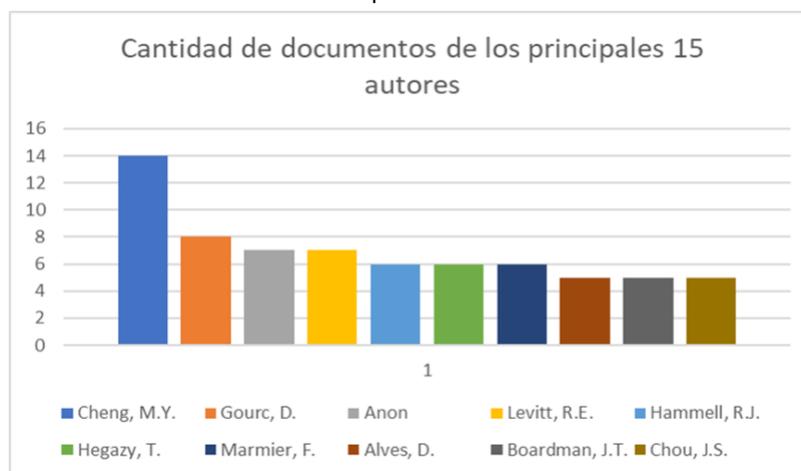
Figura 8
Países con más investigaciones en el tema



Fuente: Elaboración propia basada en los datos obtenidos de Scoopus

5.2. Autores

Figura 9
Principales autores



Fuente: Elaboración propia basada en los datos obtenidos de Scoopus

Como se ve en la figura 9, los autores que están explorando este campo se encuentran en una fase de ampliación de la información e investigación, tendencia que permite dar la guía para nuevos investigadores que se encuentren en la misma línea de acción.

5.3. Temas principales

Figura 10
Temas relacionados de los autores
de los artículos analizados

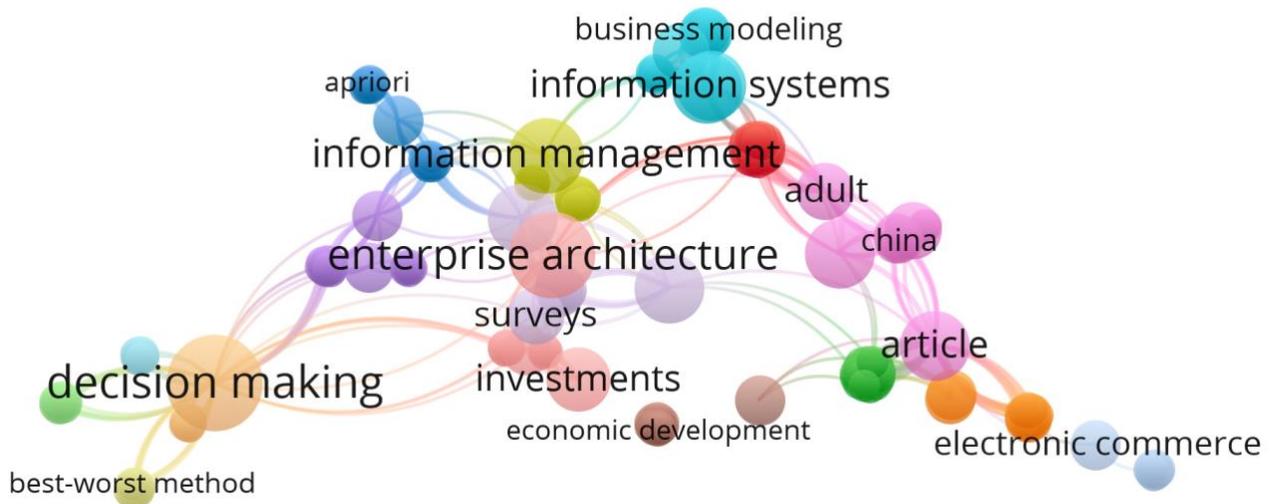


Fuente: Elaboración propia usando
VosViewer a partir de los datos de Scoopus

Los autores tienen la tendencia a desarrollar temas relacionados con la gerencia de proyectos, la arquitectura empresarial y maneras de administrar la información como se observa en la figura 10.

5.4. Palabras indexadas

Figura 11
Temas indexados en los artículos analizados



Fuente: Elaboración propia usando VosViewer
a partir de los datos de Scoopus

Como vemos en la figura 11, la ciencia de datos logra interconectar distintas áreas de conocimiento, siendo un gran soporte para la toma de decisiones organizacionales (Martin van den Berg, 2019) permitiendo asegurar el éxito financiero de las empresas, indistintamente de su tipo de carácter, desde comercio electrónico (Shweta y Deepak, 2019) hasta la arquitectura empresarial de un hospital (Ajera, 2019) o el uso de la misma como apoyo para los profesionales de la salud (Shahbaz, 2019).

6. Discusión de resultados

Dado el incremento en los últimos años de la investigación en análisis de riesgos (como se muestra en la figura 6), se ha buscado un método que permita evaluar el riesgo empresarial de manera eficiente y objetiva; para esto,

se ha decidido analizar el gran volumen de datos que cada empresa produce a diario, con el fin de encontrar patrones que permitan visualizar el comportamiento de la empresa de manera evidente (Shahnawaz Muhammed, 2018). Es aquí donde las empresas han comenzado a usar algoritmos de ciencia de datos; ya que se hace necesario determinar cómo se relacionan distintos elementos de una empresa, en un evento determinado. El inconveniente presentado; es que estos algoritmos necesitan siempre un apoyo que pueda determinar si los datos son correctos, analizar si las reglas que rigen los algoritmos están bien definidas o actualizar los datos necesarios para el buen funcionamiento de los algoritmos (Maozhu Jin, 2018).

Las técnicas de machine learning demuestran resultados positivos en análisis de riesgos financieros; específicamente, el aprendizaje profundo (Deep learning); la implementación de redes neuronales con un número alto de capas ocultas, permiten solucionar inconvenientes que presentan las técnicas más comunes del aprendizaje de máquina, facilitando evadir los tediosos pasos de la ingeniería de características (feature engineering), pues el deeplearning demuestra su capacidad para abstraer características informativas de manera autónoma; además los resultados de predicción de los modelos de redes neuronales profundas (DNN) resultan tener más exactitud respecto a los clasificadores habituales, concluyendo que las DNN son una herramienta útil para apoyar la toma de decisiones en el caso de estudio de riesgos financieros (Kim y Yang, 2019).

Se encontró que existen herramientas y metodologías que son difíciles de aplicar y costosas de implementar, requiriendo personas con conocimientos muy particulares dentro del campo de la ciencia de datos; además algunas de estas no permiten incorporar actividades de gestión de proyectos como la etapa del análisis de riesgos, (Rodríguez y Díaz, 2009) (Moine, Gordillo y Haedo, 2011). Asimismo algunas técnicas de minería de datos pueden ser aplicadas bajo los parámetros del PMBOK, dentro de las etapas de la gestión de proyectos en casos determinados (Castro et al, 2018).

En los últimos años ha crecido la producción y el desarrollo de proyectos que aplican ciencia de datos y algoritmos de uinteligencia artificial en la toma de decisiones.

7. Conculsiones

La investigación en este tema está aún en una etapa exploratoria, en donde los trabajos realizados apuntan a que existe la posibilidad de que el uso de las ciencias de datos tenga una mayor aceptación en un futuro, esperando un gran avance en el área, para que los inconvenientes presentados como la necesidad de que exista un actor que controle los algoritmos para mantenerlos o la necesidad de establecer algoritmos puntuales para cada proyecto se vean completamente superados; de tal manera que es necesario desarrollar metodologías que permitan generar simulaciones de los riesgos financieros que sean aplicables dentro de los pasos generales del análisis financiero de proyectos, permitiendo hacer uso de estas en un mayor número de situaciones.

Referencias bibliográficas

- Aalst, W. (2016). *Process Mining: Data science in action*. Berlín: Springer. doi:<https://doi.org/10.1007/978-3-662-49851-4>
- Agarwal, R., y Dhar, V. (2014). Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. *Information Systems Research*, 25(3), 443-448. doi:<https://doi.org/10.1287/isre.2014.0546>
- Aguilar, L. J. (2016). *Big Data. Análisis de grandes volúmenes de datos en organizaciones*. Alfaomega.
- Ajera, A. H. (2019). Enterprise architecture in hospitals: Resolving incongruence issues. *Studies in Health Technology and Informatics*, 659-663.

- Amaya, J. A. (2010). Toma de decisiones gerenciales. Métodos cuantitativos para la administración. ECOE Ediciones.
- Assaff, R. (2006). PMBOK – El Cuerpo de Conocimientos de la Gestión de Proyectos. En T. V. Jan Van Bon, Frameworks For IT Management (pág. 226). Van Haren Publishing.
- Barredo, A. et al. (2018). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible A. doi: <https://doi.org/10.1016/j.inffus.2019.12.012>
- Bazzani, C. y Cruz, E. (2008). Análisis de riesgos en proyectos de inversión un caso de estudio. Scientia et Technica, 1(38), 309 - 314.
- Bernstein, L. (1997). Fundamentos de análisis financieros. McGraw-Hill Interamericana de España.
- Brenes, A. I. (2016). Toma de decisiones gerenciales. Tecnología en marcha, 44-51.
- Britos, H. G. (2006). Minería de Datos. Nueva Librería.
- Brock, D. C. (2018). Learning from Artificial Intelligence's Previous Awakenings: The History of Expert Systems . AI Magazine, 39(3), 3-15. doi:<https://doi.org/10.1609/aimag.v39i3.2809>
- Bruce, P., Bruce, A. y Gedeck, p. (2020). Practical Statistics for Data Scientists: 50 essential concepts using R and Python (2ª ed.). O'Reilly Media.
- Castro, G. et al. (2018). Método para el aseguramiento de ingresos basado en análisis de riesgos y computación con palabras. Revista Iberica de Sistemas e Tecnologias de Informacao, 126 - 141.
- Cheng, M. Y., & Roy, A. F. (Agosto de 2010). Evolutionary fuzzy decision model for construction management using support vector machine. Expert Systems with Applications, 37(8), 6061-6069. doi:10.1016/j.eswa.2010.02.120
- Cheng, M. Y., Tsai, M. H., Ko, C. H., & Chen, P. H. (2005). Evolutionary fuzzy neural inference system for prediction of project success. The 9th World Multi-Conference on Systemics, Cybernetics and Informatics, (págs. 235-239). Orlando.
- Cleveland, W. (2001). Data Science: An Action Plan for Expanding the Technical Areas of the Field of Statistics. International Statistical Institute, 21-26.
- Del Carpio, J. (2006). Análisis del riesgo en la administración de proyectos de tecnología de información. Industrial Data, 8(1), 104 - 107.
- Espinoza, S. F. (2010). Los proyectos de inversión. Evaluación financiera. Costa Rica: Editorial Tecnológica de Costa Rica.
- Faulín, J. y Juan, A. (2005). Simulación de Monte Carlo con excel. Obtenido de http://juanalmendras.tripod.com/sitebuildercontent/sitebuilderfiles/simulacion_mc.pdf
- Gartner (2020). Gartner Glossary. Obtenido de <https://www.gartner.com/en/information-technology/glossary/big-data>
- Haque, A., Ayyar, A. V. y Singh, S. (2018). A meta data mining framework for botnet analysis. International Journal of Computers and Applications, 41(5), 392-399.
- Hipodec, R. (2018). ¿Qué es el riesgo financiero? Universidad Panamericana.
- IntelliPaat. (s.f.). <https://intellipaata.com/blog/what-is-data-science/>.

- Islam, K. y O’Gorman, M. (2019). Microcredit contract design: A macroeconomic evaluation. *World Development*, 124(1).
- Kenneth Cukier, V. M. (2013). *Big data : la revolución de los datos masivos*. Turner Publicaciones S.L.
- Kim, A., y Yang, Y. (2019). Deep Learning Predict Risky Retail Investors? A Case Study in Financial Risk Behavior Forecasting. *European Journal of Operational Research*, 1-37.
doi:<https://doi.org/10.1016/j.ejor.2019.11.007>
- Koyuncugil, A. S., & OZgulbas, N. (2012). Financial early warning system model and data mining application for risk detection. *Expert Systems With Applications*, 39(6), 6238-6253.
doi:<https://doi.org/10.1016/j.eswa.2011.12.021>
- Lledó, P. y Rivarola G. (2007). *Gestión de proyectos: Cómo dirigir proyectos exitosos, coordinar los recursos humanos y administrar los riesgos*. Buenos Aires: Pearson.
- Lengauer, T. (2020). Statistical Data Analysis in the Era of Big Dat. *Chemie Ingenieur Technik*, 92(7), 187–206.
doi:[10.1002/cite.202000024](https://doi.org/10.1002/cite.202000024)
- Maozhu Jin, Y. W. (2018). Application of Data Mining Technology in Financial Risk Analysis. *Wireless Pers Commun*, 3699–371.
- Muñoz, M., Arenas, L. y García, C. (2017). Riesgo financiero: Una aproximación cualitativa al interior de las mipymes en Colombia. *Aglala*, 8(1), 139 - 160.
- Menichini, A. (2004). Value at Risk : metodología de administración del riesgo financiero. *Invenio*, 127-138.
- Miranda, J. J. (2005). *Gestión de proyectos. Identificación, Formulación y evaluación*. MM Editores.
- Moine, J. M., Gordillo, S. y Haedo, A. (2011). Análisis comparativo de metodologías para la gestión de proyectos de minería de datos. *Congreso Argentino de Ciencias de la Computación*, 17, 931-938.
- Mora, J. P. (s.f.). Qué es la Ciencia de Datos, el aprendizaje automático(ML), el Big Data y cuales son sus usos? Obtenido de <https://www.javeriana.edu.co/documents/12847/10949798/Qu%C3%A9+es+la+ciencia+de+datos/aa1e64ec-0961-4ef9-a4e5-472fe5774c0b>
- Padilla, M. C. (2012). *Gestión financiera*. ECOE Ediciones.
- Pérez, C. y Santín, D. (2008). *Minería de datos: técnicas y herramientas*. Madrid: Paraninfo.
- Provost, F. y Fawcett, T. (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. Estados Unidos: O'Reilly Media.
- Mukherjee, R. y Kar, P. (2017). A Comparative Review of Data Warehousing ETL Tools with New Trends and Industry Insight. 7th International Advance Computing Conference (IACC).
- Rodríguez, Y. Díaz, A. (2009). Herramientas de minería de datos. *Revista Cubana de Ciencias Informáticas*, 3(3-4), 73 - 80.
- Russel, S. y Norving, P. (2008). *Inteligencia artificial. Un enfoque moderno*. México: Pearson.
- Russell Kelly, A. T. (s.f.). Jet-Software. Obtenido de <https://www.jet-software.com/en/data-integration-fast-and-easy-iri-etl-tools/etl-diagram/>
- Saltz, J. S., y Shamshurin, I. (2016). Big data team process methodologies: A literature review and the identification of key factors for a project's success. 2016 IEEE International Conference on Big Data, (págs. 2872-2879). Washington. doi:[10.1109/BigData.2016.7840936](https://doi.org/10.1109/BigData.2016.7840936)

- Saltz, J. S., Shamshurin, I., y Connors, C. (2017). A framework for describing big data projects. En Lecture Notes in Business Information Processing (págs. 183-195). doi:10.1007/978-3-319-52464-1_17
- Shahbaz, M. G. (2019). Investigating the adoption of big data analytics in healthcare: the moderating role of resistance to change. *Journal of Big Data*.
- Shahnawaz, M. y Kuklakli, A. (2018). Big Data in Mitigating Systemic Supply Chain Risks: A Systems Perspective with Case Studies. *Journal of Business Research-Turk*, 15-32. doi:10.20491/isarder.2018.460
- Shweta, P. y Deepak, C. (2019). Engaging m-commerce adopters in india: Exploring the two ends of the adoption continuum across four m-commerce categories. *Journal of Enterprise Information Management*, 32(1), 191 - 210.
- Van den Berg, M., Slot, R., van Steenbergen, M., Faasse, P. y van Vliet, H. (2019). How Enterprise Architecture Improves the Quality of IT Investment Decisions. *Journal of Systems and Software*, 152, 135 - 150.
- Wu, J. (s.f.). Statistics = Data Science. Obtenido de <http://www2.isye.gatech.edu/~jeffwu/presentations/datascience.pdf>.

Esta obra está bajo una Licencia Creative Commons
Atribución-NoCommercial 4.0 International

