

Big Data analysis in Machine Learning for strategic planning of resilient supply chains

Análisis del Big Data en Machine Learning para la planificación estratégica de las cadenas de suministro resilientes

Alondra Guadalupe, MORA-HERNÁNDEZ^{1*}

¹*Instituto Tecnológico de Ciudad Cuauhtémoc, Technological Av. 137, Ciudad Cuauhtémoc, Chihuahua, C.P. 31500, México. (0009-0003-3884-0187)*

Sent date: 01/August/2023 Acceptance date: 30/October/2023

Abstract:

This paper aims to describe the temporal changes in the processing of information from the simple compilation in the past to require a detailed categorization and analysis today as a method of support in decision making. The phenomenon that is developed below is the examination of Big Data in which, for a better understanding, some of the basic concepts that integrate it and its operability are exposed, as well as some of the analytical techniques commonly used according to the purpose pursued. Likewise, the usefulness of the applications in everyday life is highlighted, emphasizing in the organizational environment, mainly in supply chains, the ability to respond quickly to disruptive events, since the constantly changing environment forces companies to implement strategies that allow an immediate response to risks and opportunities through the use of contemporary technologies that reduce the adverse effects. By detailing key concepts, understanding how Big Data works in Machine Learning and applications, the structure of a model for supply chains is detailed, which will allow us to consider the importance of adopting these practices.

Keywords: Big Data, Machine Learning, Resilient Supply Chain

Resumen:

Este trabajo pretende describir los cambios temporales en el procesamiento de la información pasando de la simple compilación antiguamente a requerir una detallada categorización y análisis en la actualidad como un método de sustento en la toma de decisiones. El fenómeno que se desarrolla a continuación es el examen del Big Data en el cual para un mejor entendimiento se exponen algunos de los conceptos básicos que lo integran y la operatividad del mismo, también se detallan algunas de las técnicas analíticas comúnmente empleadas según el fin perseguido, del mismo modo se resalta la utilidad de las aplicaciones en la cotidianidad enfatizando en el ambiente organizacional principalmente en las cadenas de suministro, aquella capacidad de responder rápidamente a sucesos disruptivos, ya que el entorno en constante cambio obliga a las empresas a implementar estrategias que permitan una respuesta inmediata a los riesgos y oportunidades mediante el uso de tecnologías contemporáneas que reduzcan los efectos adversos. Por medio de la pormenorización de los

conceptos clave, el entendimiento del funcionamiento del Big Data en Machine Learning y de las aplicaciones se detalla la estructura de un modelo para las cadenas de suministro que nos permitirá considerar la importancia de adoptar estas prácticas.

Palabras clave: Big Data, Machine Learning, Cadena de Suministro Resiliente

*Corresponding autor. E-mail: alondramora2611@gmail.com,
Tel. 523318339990

1. Introducción

Desde tiempos antiguos las sociedades han establecido una administración sobre el manejo de su información con el fin de registrarla, procesarla y utilizarla para beneficio propio, misma que plasman principalmente en instrumentos manuales y codificados. Sin embargo, en épocas recientes se han modificado las herramientas para la gestión de estas referencias, trayendo consigo nuevas metodologías y conceptos a ser objeto de observación, máxime a partir de la instauración y evolución de la Revolución Digital donde se popularizaron los estudios sobre nuevas definiciones y teorías que han surgido con este suceso, tal es el caso del concepto de sociedad de la información, mismo que define cuán imprescindible es la divulgación en los vínculos interpersonales y en todas las acciones que realizamos día con día incluyendo las decisiones que tomamos en el entorno empresarial.

Actualmente la mera recopilación de contenidos por parte de las compañías no es suficiente ni aporta utilidad por sí misma, es indispensable emplear herramientas en el Big Data para procesar dichos informes y ejecutar un análisis que permita reconocer posibles problemáticas y formular resoluciones, asimismo encontrar oportunidades mediante la exploración de tendencias que permitan añadir valor por medio de la implementación de mejoras en los procesos, disminución de costos, esquematización de nuevos productos y servicios, de igual forma reconocer el mercado clave para un correcto enfoque del marketing, por mencionar algunos. El análisis predictivo de instrumentos como Machine Learning, facilitan dicha interpretación no sólo en el sector corporativo sino en los ramos de salud, farmacéutico, automotriz, aeroespacial y de manufactura, es por ello que el objetivo de este trabajo es presentar un análisis y la estructura de un modelo de planificación en la cadena de suministro resiliente con el fin de considerar la importancia en su adopción.

2. Materiales y Métodos

2.1 Procesamiento de Big Data en Machine Learning

El Big Data es un fenómeno donde las empresas y organizaciones aprenden a administrar el crecimiento de sus datos, (Sansores, 2020), consiste en la compilación de enormes cantidades de información que cumplen ciertas características como el volumen, la velocidad con la que se generan y la variedad de contenido (denominadas las 3 V), y es debido a la magnitud y complejidad de la información que tienen las compañías en el presente que se demandan

centros de almacenamiento nombrados Data Warehouses donde se categorizan de acuerdo a la naturaleza de los informes, ya sean estructurados, no estructurados o semiestructurados y permiten realizar consultas con base al conocimiento obtenido de un amplio repertorio de fuentes. De acuerdo a Cisco, Big Data es un gran negocio debido a la importancia de los volúmenes de referencias al permitir a las entidades ser más predictivas y ágiles ante un escenario cambiante. No obstante, pese a tener disponible esta información si no se elabora un profundo análisis partiendo del objetivo que se quiere obtener resulta inservible dicha recopilación.

Las nuevas tecnologías aplicadas al Big Data, por ejemplo, la Inteligencia Artificial (IA), han traído consigo nuevos métodos más simplificados que funcionan como técnicas analíticas tales como: minería de datos, visualización, análisis estadístico y aprendizaje automático o Machine Learning (Ahmed Osussous, 2018). Considerando el resultado buscado para la toma de decisiones en la planificación de cadenas de suministro, Machine Learning resulta la práctica más adecuada para nuestro objeto de análisis. Consiste en un software que origina autónomamente conjuntos de reglas sustentadas en sucesos de reconocimiento instantáneo, es decir el estudio de modelos complejos y produce un análisis predictivo refiriendo variables de acontecimientos pasados y de la correlación entre los mismos que faculte pronosticar el comportamiento futuro en situaciones precisas. Los algoritmos bajo los cuales funciona Machine Learning están configurados para clasificar la información, ubicar patrones de relación, esbozar resultados y a partir de ello tomar decisiones fundamentadas, dichos algoritmos tienen la facilidad de poder combinarse para obtener una mayor precisión en el resultado (Figura 1).

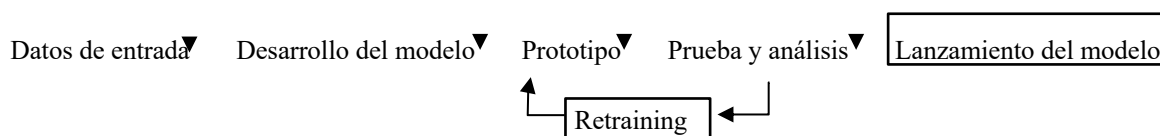


Figura 1. Funcionamiento del proceso de Machine Learning

Estas tácticas algorítmicas se pueden emplear en los cuatro modelos de aprendizaje que se utilizan en Machine Learning cuya elección depende de la procedencia de los datos y el resultado esperado: supervisado, no supervisado, semisupervisado o de refuerzo.

El éxito crucial al procesar los datos del Big Data en Machine Learning reside en la introducción de las referencias en los algoritmos de aprendizaje, es decir, a mayor cantidad de información es más arduo y completo el algoritmo o la instrucción insertada por lo que se trata de una solución Ad Hoc a las necesidades de la compañía lo cual difiere de otros algoritmos donde primero se crea la instrucción y después se introducen las notas.

3. Resultados y Discusión

3.1 Aplicaciones del análisis predictivo en Machine Learning

La analítica avanzada del Big Data ha eficientizado el procesamiento de la información mediante tres sistemas: (Tabla 1).

Tabla 1. Tipos de sistemas para procesamiento de información

TIPO DE ANALÍTICA	USO
Descriptiva	Permite concentrar el Big Data en datos más pequeños para hacerlos más manejables.
Predictiva mediante	Se emplea para pronosticar posibles acontecimientos la minería de datos y patrones de comportamiento.
Prescriptiva	Para esta etapa se requiere ejecutar la dos previamente como soporte para aplicar la prescriptiva en la toma de decisiones.

El análisis predictivo es una de aplicaciones más adoptadas por las organizaciones ya que permite optimizar y admitir una segmentación compleja para automatizar los datos y tener una mejora en las operaciones a fin de tener una ventaja sobre la competencia mediante el uso de Machine Learning (Watson, 2014), el cual es gestionado en las corporaciones habitualmente por los departamentos de Business Intelligence o Advanced Analytics quienes desarrollan este sistema a través de medios propios o mediante tercerización del servicio con compañías especializadas como IBM, SAS, Oracle o Microsoft. Estas firmas comparten cinco áreas de aplicación en la analítica descriptiva que proporcionan y favorecen la rentabilidad de las entidades: segmentación de clientes, personalización de la oferta, evaluación de riesgo del abandono de la relación, éxito de una publicidad y tasa de deserción de los clientes, lo cual permite a las empresas realizar valoraciones de distintos escenarios de manera inmediata.

En el tiempo actual algunos de los ejemplos más comunes del uso de esta herramienta son los ERP con la automatización de procesos en los cuales las bases de datos que integran contienen data sets amplios y dispares acerca de estadísticas de ventas, evaluación de consumidores, tendencias de mercado y registros de gestión en cadenas de suministro, así los algoritmos de Machine Learning pueden encontrar correlaciones entre estas referencias ofreciendo información concentrada y precisa a todas las áreas de un negocio para la toma de resoluciones.

En las cadenas de suministro se puede resaltar la eficacia de este instrumento con el mantenimiento predictivo, consiste en detectar posibles fallas e ineficiencias en los mecanismos que pueden ocasionar costos y interrupciones en los eslabones mediante la instalación en los dispositivos y máquinas de IoT sensores de enlace que coadyuvan en la visualización de estas situaciones. Asimismo, en la comercialización de mercancías o

materias primas se observa la empleabilidad al obtener conocimiento sobre los flujos de datos en relación a los precios de acuerdo a históricos sobre condiciones del mercado, estadísticas, regiones, etcétera.

4. Modelo de aprendizaje en las cadenas de suministro

De acuerdo a un estudio realizado por Oxford Economics durante el primer trimestre del 2023 donde se entrevistó a mil ejecutivos de planificación de cadenas de suministro, se detectaron las siguientes áreas de oportunidad para lograr una mayor resiliencia en los procesos de las cadenas de suministro. (Figura 2).

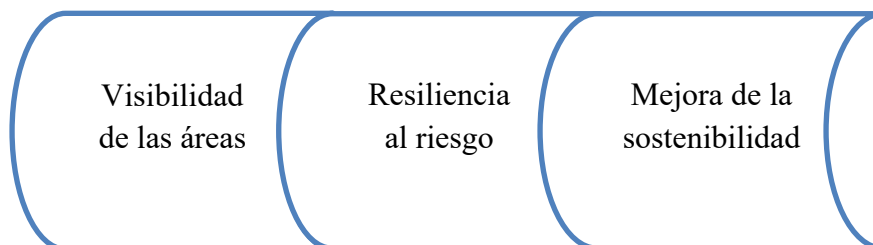


Figura 2. Áreas de oportunidad en las cadenas de suministro.

En nuestros días la gestión de la cadena de suministro resulta más compleja debido a la variedad de opciones derivada de las variables en constante cambio que aumentan la incertidumbre en la toma de decisiones, los nuevos enfoques se basan principalmente en el cambio en la planificación tradicional de la demanda y la reacción ante contingencias inesperadas utilizando modelos más dinámicos y predictivos

La determinación de correlaciones entre las variables y patrones de comportamiento en un conjunto de informes mediante el análisis predictivo es posible gracias a la disponibilidad de una gran cantidad de contenido sometida a algún proceso de aprendizaje delimitado en los cuatro modelos de Machine Learning: supervisado, no supervisado, semisupervisado o de refuerzo. Considerando las características que poseen cada uno, se opta porque el modelo de aprendizaje no supervisado es el más adecuado según nuestro objetivo de analizar la información que funja como soporte para una planeación estratégica en las cadenas de suministro resilientes, dado que en este modelo se estudian las referencias de entrada identificando la asociación entre variables empleando los datos relevantes y accesibles. El análisis de secuencias genéticas, investigaciones de mercado y ciberseguridad son algunos ejemplos del uso de este modelo el cual incluye tres categorías de algoritmos para su funcionamiento (Tabla 2).

Tabla 2. Algoritmos del aprendizaje no supervisado.

ALGORITMO	DESCRIPCIÓN	MODELOS	REPRESENTACIÓN
Clúster o agrupación	Agrupación de objetos basados en sus características	Modelo partitioning clustering	K-Medias $deuc(p,q) = (xp-xq)^2 + (yp-yq)^2$
		Modelo jerárquico	Estructura de árbol Dendograma
Reducción de dimensionalidad	Busca reducir el número de variables en un conjunto de datos.	Técnicas de selección de variables	Correlación o consistencia
		Análisis de componentes principales	Varianza
Reglas de asociación	Relaciones de asociación	Soporte	Moda
		Confianza	Probabilidad de que ocurra B si ocurre A (A B)
		Lift de la regla	Nivel de dependencia Lift<1 Lift=1 Lift>1

La optimización de este modelo depende de un exhaustivo análisis de relevancia que principia con el método de selección de variables para lo cual se identifican dos posibles procedimientos matemáticos: la correlación que permite medir el nivel de asociación descartando aquellas variables que tengan valor negativo, igual a la variable principal o próxima a cero y el criterio de consistencia que permite eliminar las variables que afecten el resultado deseado.

5. Conclusiones

El escenario cambiante en la que nos desenvolvemos de forma cotidiana nos obliga a ir adoptando nuevas tecnologías que coadyuven a entender los cambios acontecidos y nos permitan estar preparados ante nuevos retos en distintos contextos, el ámbito empresarial es uno de los campos que más lentitud ha presentado en esta adaptación ocasionando pérdidas económicas relevantes al mostrar una fuerte interrelación entre las actividades de los departamentos que la integran, una de las más significativas es la cadena de suministro. Las cadenas de suministro que habían prevalecido durante los últimos 30 años han tenido que reconstruirse tras el escenario ocasionado por la pandemia por Covid-19, dando preeminencia a la gestión de riesgos y no sólo a la eficiencia en la producción evaluando la

gestión de costos y tiempos en cada nodo del eslabón. Los pronósticos tradicionales basados en datos históricos han resultado ineficaces en un entorno de constantes eventualidades por lo que los modelos actuales requieren una percepción más precisa y expedita que permita definir estrategias de acuerdo a los recursos disponibles y a la predicción de múltiples condiciones.

El uso de tecnologías como el análisis predictivo de Big Data en Machine Learning proporciona una evaluación inmediata de los posibles impactos en la operación facilitando este proceso y reduciendo los tiempos de respuesta. Ahora bien, la pesquisa realizada muestra la relevancia de la selección de información y variables en la utilización de esta herramienta en la búsqueda del objetivo deseado, en otras palabras, pese a que hoy en día contemos con instrumentos tecnológicos sofisticados que faciliten la ejecución de actividades, resulta trascendental el criterio humano en este desarrollo.

6. Nomenclatura

- **Advanced Analytics:** estudio de datos mediante herramientas sofisticadas y técnicas analíticas avanzadas como minería de datos, aprendizaje automático, coincidencia de patrones, pronósticos, visualización, análisis de redes y clústeres, estadísticas multivariadas y procesamiento de redes neuronales.
- **Big Data:** designado también como datos masivos, se refiere a un gran volumen de datos que crece de forma exponencial.
- **Business Intelligence (BI):** software alimentado por datos de negocios que presenta la información concentrada en gráficos para ayudar a la toma de decisiones.
- **Data Warehouse:** sistema de gestión de datos diseñado para centralizar grandes cantidades de datos de múltiples fuentes.
- **ERP (Enterprise Resources Planning):** sistema que ayuda a automatizar los procesos empresariales de distintas áreas como finanzas, ventas, cadena de suministro, recursos humanos y operaciones.
- **Inteligencia Artificial (AI):** habilidad de los ordenadores para realizar actividades que normalmente requieren inteligencia humana.

7. Referencias

- Ahmed Oussous, F. (2018). Big Data Technologies: a survey. *Journal of King Saud University- Computer and Information Sciences*.
- Escuela de Postgrado de la Universidad Católica San Pablo. (2023). Big data: definición, tipos, características y beneficios. Available at: <https://postgrado.ucsp.edu.pe/articulos/que-es-big-data/>. Accessed: August 20th, 2023.

- Espino, C. (2017). Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo- herramientas Open Source que permiten su uso. Trabajo de fin de grado, Universitat Oberta de Catalunya, Barcelona, España.
- González-Iglesias, R. (2015). Herramientas para el BigData y Machine Learning. Trabajo de fin de Grado en Estadística, Facultad de Ciencias, Universidad de Valladolid, España.
- IBM. (2023). ¿Qué es Business Intelligence?. Available at: <https://www.ibm.com/mx-es/topics/business-intelligence>. Accessed: August 20th, 2023.
- IBM (2023). ¿Qué es Machine Learning?. Available at: <https://www.ibm.com/mx-es/analytics/machine-learning>. Accessed: August 18th, 2023.
- Information Technology Gartner Glossary. (2023). Advanced Analytics. Available at: <https://www.gartner.com/en/information-technology/glossary/advanced-analytics>. Accessed: August 20th, 2023.
- Kacfeh Emani, C., Cullot, N. & Nic, C. (2015). *Comprehensible Big Data: una encuesta*. ELSEVIER.
- Martín-Romero, A. (2020). Big Data. Técnicas de machine learning para la creación de modelos predictivos para empresas. Trabajo fin de carrera, Facultad Ciencias Económicas y Empresariales, Universidad Pontificia, Madrid, España
- Oracle. (2023). ¿Qué es un almacén de datos?. Available at: <https://www.oracle.com/mx/database/what-is-a-data-warehouse/>. Accessed: August 19th, 2023.
- Oxford Economics. (2023). Construir visibilidad con planificación de la cadena de suministro. *Oxford Economics*.
https://www.sap.com/latinamerica/products/scm/integrated-business-planning.html?pdf-asset=54f5ef1e-7e7e-0010-bca6-c68f7e60039b&page=1&ut=L2xhdGluYW1lcmljYS9kb2N1bWVudHMvMjAyMy8wNy81NGY1ZWYxZS03ZTdlLTAwMTAtYmNhNi1jNjhmN2U2MDAzOWI%3D&client_id=c5cc1c92-3bbd-11ee-9161-29f9e888fde7
- Roa, S. & Loaiza, H. (2011). Evaluación de técnicas para el análisis de relevancia basadas en filtros sobre imágenes radiológicas. En *Revista Ingeniería Biomédica*.
- Rouhiainen, L. (2018). *Inteligencia Artificial. 101 cosas que debes saber hoy sobre nuestro futuro*. Editorial Alienta, España.
- Sansores, J. (2020). Big Data: Conceptos Básicos, Tecnologías y Aplicaciones. Trabajo monográfico para obtener el grado. División de Ciencias e Ingeniería, Universidad de Quintana Roo.
- SAP (2023). Key Supply Chain Issues, Explained. Available at: <https://www.sap.com/latinamerica/insights/viewpoints/key-supply-chain-issues-tech-strategies.html>. Accessed: August 19th, 2023.
- SAP (2023). ¿Qué es machine learning?. Available at: <https://www.sap.com/latinamerica/products/artificial-intelligence/what-is-machine-learning.html>. Accessed: August 18th, 2023.

- SAP (2023). ¿Qué es una cadena de suministro resiliente? Available at: <https://www.sap.com/latinamerica/products/scm/integrated-business-planning/what-is-a-resilient-supply-chain.html#:~:text=Las%20cadenas%20de%20suministro%20m%C3%A1s,que%20surja%20en%20el%20futuro>. Accessed: August 18th, 2023.
- Watson, H. (2014). Big Data Analytics: conceptos, tecnologías y aplicaciones. En *Comunicaciones de la Asociación de Sistemas de Información*.