



**UNIVERSIDAD POLITÉCNICA SALESIANA SEDE
GUAYAQUIL CARRERA DE INGENIERÍA DE
SISTEMAS**

**RIESGO EN EL USO DE MODELOS DE INTELIGENCIA ARTIFICIAL.
ANÁLISIS BASADO EN MODELO CAUSALES**

Trabajo de titulación previo a la obtención del
Título de Ingeniero de Sistemas

AUTOR 1: CRISTOPHER ALEJANDRO RAMIREZ MATEO

AUTOR 2: JOSEPH JORDAN ALCIVAR PINEDA

TUTOR: MIGUEL ÁNGEL QUIROZ MARTÍNEZ

Guayaquil – Ecuador

2022

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Joseph Jordan Alcivar Pineda con documento de identificación N° 0929242816 manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 08 de febrero del año 2022

Atentamente,



Joseph Jordan Alcivar Pineda 0929242816

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Joseph Jordan Alcivar Pineda con documento de identificación No. 0929242816, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor(a) del Artículo Académico: “Riesgo en el uso de modelos de inteligencia artificial. Análisis basado en modelos causales”, el cual ha sido desarrollado para optar por el título de: Ingeniero de Sistemas, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 08 de febrero del año 2022

Atentamente,



Joseph Jordan Alcivar Pineda
0929242816

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Miguel Ángel Quiroz Martínez con documento de identificación N° 0922799655, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: RIESGOS EN EL USO DE MODELOS DE INTELIGENCIA ARTIFICIAL. ANÁLISIS BASADO EN MODELOS CAUSALES, realizado por Joseph Jordan Alcivar Pineda con documento de identificación N° 0929242816, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 08 de febrero del año 2022

Atentamente,



Firmado electrónicamente por:
**MIGUEL ANGEL
QUIROZ
MARTINEZ**

Miguel Ángel Quiroz Martínez

0922799655

DEDICATORIA

El presente trabajo está dedicado a mi familia por haber sido mi apoyo a lo largo de toda mi carrera universitaria y a lo largo de mi vida. A todas las personas especiales que me acompañaron en esta etapa, aportando a mi formación tanto profesional y como ser humano.

AGRADECIMIENTO

Agradezco a mi madre por toda su ayuda y motivación para impulsarme siempre a seguir adelante y nunca dejar que me rinda sin importar que tan difíciles sea la situación, agradezco a mi tía Teresa que es como mi segunda madre por siempre estar ahí dándome fuerzas y consejos a lo largo de toda mi carrera, agradezco a mis profesores y a todas las personas que he conocido a lo largo de este camino que me ha llevado a este momento tan importante en mi vida.

RESUMEN

En los diferentes contextos en donde se desenvuelve el ser humano se ha podido observar que, con el uso masivo de plataformas digitales, el rol del análisis de datos ha aumentado, lo que ha producido un incremento del interés de usarlos por medio de los navegadores. Bajo la inteligencia artificial, tanto en el uso de software, el almacenamiento de información y el paralelismo con las diferentes redes sociales, han generado un producto de alto poder predictivo, con interpretación automática y semántica según los datos ingresados. Este estudio fue realizado con un enfoque cuantitativo, con criterio descriptivo y ha considerado el modelado del aprendizaje, el cual ha permitido adelantarse a la toma de decisiones y ha contribuido a la predicción educativa. Se concluyó que la importancia de utilizar el aprendizaje profundo y redes neuronales útiles y efectivas puede beneficiar a las empresas, por esto, se proyectan su uso en usuarios promedio, pero, mientras más se demanda de este tipo de tecnología, también pueden aparecer riesgos colaterales.

Palabras claves: Fuzzy AHP TOPSIS; Aprendizaje Profundo; Navegadores.

ABSTRACT

In the different contexts in which the human being develops, it has been observed that, with the massive use of digital platforms, the role of data analysis has increased, which has produced an increased interest in using them through browsers. Under artificial intelligence, both in the use of software, the storage of information and the parallelism with the different social networks, they have generated a product of high predictive power, with automatic and semantic interpretation according to the data entered. This study was conducted with a quantitative approach, with descriptive criteria and has considered the modeling of learning, which has allowed us to anticipate decision making and has contributed to educational prediction. The objective was to establish the risks of using fuzzy AHP TOPSIS models for deep learning in browsers. It was concluded that the importance of using deep learning and useful and effective neural networks can benefit companies, for this reason, their use is projected in average users, but the more demand for this type of technology, collateral risks may also appear.

Key words: *Fuzzy AHP TOPSIS; deep learning; browsers.*

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN.....	10
2. MATERIALES Y METODOS.....	10
2.1 Materiales.....	10
2.2 Recreación de una neurona	11
3. METODOLOGIA	12
3.1 Riesgos de la Inteligencia artificial	13
4. RESULTADOS.....	14
5. CONCLUSIÓN.....	16
REFERENCIAS	17

1. INTRODUCCIÓN

En todos los ámbitos en donde se desenvuelve el ser humano se están involucrando aspectos tecnológicos de vanguardia, de tal manera que se está volviendo algo imprescindible.

Desde el siglo XX, Santiago Ramón y Cajal se anticiparon en los estudios neurocientíficos sobre las células ramificadas del cerebro, por lo cual se le fue otorgado el premio Nobel en el año 1906. Después de un siglo la tecnología se expande y surge el aprendizaje automático. Con la llegada de la Inteligencia Artificial (IA), ha habido un considerable avance tecnológico, pero a más de oportunidades, también se han observado varios riesgos, que derivan en accidentes o resultados no esperados, por el mal uso de los mismos.

Esta investigación tiene el objetivo de analizar modelos de la IA al establecer los riesgos del uso de modelos fuzzy AHP TOPSIS para deep learning en navegadores para el aprendizaje profundo. Este tipo de tecnología a menudo genera, preserva y comparte información privada (Onik, 2019). Cada vez un mayor número de personas se convierten en clientes mediáticos de los contenidos, junto con la exposición de la información que atraen a las compañías y demás instituciones deseosas de captar datos por medio de la IA. Toda una serie de recursos son usados con los diferentes métodos y técnicas algorítmicas para el aprendizaje automático.

2. MATERIALES Y METODOS

2.1 Materiales

En la actualidad se está utilizando este término, puesto que es la era de la tecnología, con el uso de la IA, que está impactando al mundo, pero no se puede hablar de los factores favorables sino también de los riesgos que son producidos.

La IA se ejecuta por medio de máquinas y para esto se establecen las tecnologías de la computación de última generación, utilizando equipos con programas de uso flexible, que permiten ejecutar acciones, que antes tomaban mucho tiempo y no se podían utilizar muchas herramientas. Dichas máquinas optimizaron tiempo y recursos, permiten realizar actividades que generan beneficios económicos, así mismo se pueden realizar otras actividades, como

calcular, jugar, comunicarse con otros usuarios, siendo ahora una necesidad general, tanto a nivel personal, como institucional.

2.2 Recreación de una neurona

Una de los retos que emergen cuando se trata de neuronas, es la posibilidad de imitar las capacidades naturales de los humanos en su funcionamiento cerebral, convirtiéndolo en un desafío extraordinario, alcanzable, escalable; este proceso de aprendizaje neuronal tiene complejidad en su recreación, donde el cerebro es considerado el mecanismo más complejo del universo. “Los nodos son conceptos causales y pueden modelar eventos, acciones, valores, metas o procesos” (Fig. 1- 2) (Leyva – Vázquez, 2013).

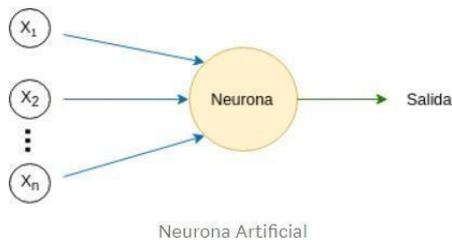


Figura 1. - Neurona artificial

$$\sum_{i=0}^m w_i x_i \quad (1)$$

Fórmula #1. Entrenamiento de Sinapsis

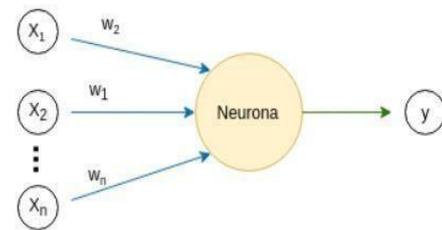


Figura 2. Conexión Neuronal Sinapsis o Pesos

$$\varphi\left(\sum_{i=0}^m w_i x_i\right) \quad (2)$$

Fórmula #2. Función de activación

La inteligencia artificial (IA), el aprendizaje automático y el aprendizaje profundo, comparten características; la IA abarca las diferentes formas en la que una máquina interactúa con el mundo que lo rodea de manera física o sonora imitando las técnicas que permiten a las computadoras parecerse a los humanos; el Machine learning es un subconjunto de inteligencia artificial que utiliza métodos estadísticos y algoritmos para permitir que las máquinas mejoren las experiencias de aprendizaje el aprendizaje automático es parte de la IA pero se les da una gran cantidad de datos para que los codifica y encuentre patrones y realice una predicción basándose en árboles de decisión. “El lenguaje profundo, es desarrollado en el área neuronal,

lo que permite imitar la conectividad del ser humano, haciendo factible el cálculo de la red neuronal multicapa” (Patel, 2017).

3. METODOLOGIA

Se emplearon Mapas Cognitivos al dominio difuso para representar los riesgos “Permiten una representación de forma más realista del conocimiento en ciclos, la vaguedad y la ambigüedad; además, un mayor nivel de usabilidad para obtener conocimiento de los expertos” (Almeida, 2019). Una ventaja radica en la probabilidad de usar el concepto de variable lingüística que aumentaría la capacidad interpretativa. “Los modelos mentales obtenidos son más cercanos al modo de pensar de los decisores, permitiendo influir en el cambio y el logro de consenso de los involucrados” (Scott R. J., Cavana R. Y., & Cameron D., 2016). En esta investigación se propone el siguiente marco de trabajo:

1. La selección de los nodos. Que corresponden a los riesgos
2. Modelar la causalidad comenzando de que existen tres posibles tipos de relaciones causales entre conceptos:

La causalidad positiva ($W_{ij} > 0$): Indica una causalidad de tipo positiva entre conceptos C_i y C_j , es decir, el incremento o disminución en el valor de C_i lleva al incremento o disminución en el valor de C_j .

La causalidad negativa ($W_{ij} < 0$): Indica una causalidad de tipo negativa entre conceptos C_i y C_j , es decir, el incremento o disminución en el valor de C_i lleva la disminución o incremento en el valor de C_j .

Inexistencia de relaciones ($W_{ij} = 0$): Indica que no existencia de relación causal entre C_i y C_j .

3. “Elaborar la matriz de adyacencia construida a partir de los valores asignados a los arcos generalmente de forma numérica” (Zhi-Qiang L., 2001) utilizando la ecuación (1). “Cuando participa un conjunto de participantes (k), la matriz de adyacencia es calculada través de un operador de agregación, como es por ejemplo la media aritmética. El método más simple consiste en encontrar la media aritmética de cada una de las conexiones para cada experto. Para k expertos, la matriz de adyacencia del MCB final (E) es obtenida como” (Bart k. & Burgess J., 2002):

$$(E^1 + E^2 + \dots + E^k)$$

$$E = \frac{\quad}{k} \quad (1)$$

4. En el siguiente análisis estático de las medidas se calculan para los valores absolutos de la matriz de adyacencia:

Outdegree, denotado por $od(v_i)$, es la suma por cada fila de los valores absolutos de una variable de la matriz de adyacencia difusa. Es una medida de la fuerza acumulada de las conexiones existentes en la variable.

Indegree, denotado por $id(v_i)$, es la suma por cada columna de los valores absolutos de una variable de la matriz de adyacencia difusa. “Mide la fuerza acumulada de entrada de la variable” (M. Leyva Vázquez & Smarandache, 2021).

La centralidad o grado total de la variable es la suma de $od(v_i)$, con $id(v_i)$, como se indica en la ecuación (2).

$$td(v_i) = od(v_i) + id(v_i) \quad (2)$$

“Finalmente, las variables se clasifican según el criterio siguiente” (Smarandache & Leyva Vázquez, 2021):

- a) Las siguientes variables son las transmisoras: $od(v_i) > 0$ e $id(v_i) = 0$.
- b) Las siguientes variables son los receptores: $od(v_i) = 0$ y $id(v_i) > 0$.
- c) Las variables ordinarias satisfacen a su vez: $od(v_i) \neq 0$ y $id(v_i) \neq 0$. Acorde a su grado de centralidad se ordenan de forma ascendente.

3.1 Riesgos en la Inteligencia artificial

Cuando una empresa desarrolla una aplicación tecnológica relacionada con la IA, debería conocer los riesgos que pueden ocasionarse y transmitirlo a los clientes que las usarán, debido a que esta labor ética evitará que sean perjudicadas las personas o empresas involucradas. “Considerando que los avances tecnológicos traen extraordinarios beneficios, pero representan mayores niveles de riesgo de modo significativo sin precedentes” (Basallo, Y. A., Senti, V. E., & Sánchez, N. M., 2018).

Empero también, quienes adquieren los productos tecnológicos deben reconocer los riesgos probables a los que se someterán con las aplicaciones utilizadas en sus negocios. De esta manera se puede anticipar y minimizar el daño que podría producirse, con acciones efectivas, priorizando los riesgos más críticos y reforzando controles, por medio de medidas de seguridad apropiadas.

Muchos riesgos más conocidos y producidos por la IA pueden anticiparse, pero pueden aparecer algunos que no se habían considerado. “Entonces es probable que las empresas no tengan en claro que hay situaciones que no podrían ser prevenidas, debido a la poca claridad de los mismos” (Liu Y. & Tang P., 2019). Un ejemplo claro es el uso de las “nubes” que guardan información importante y queda disponible a los usuarios que la necesiten. Pero también hay información privada que podría ser encontrada por los hackers y utilizada a su arbitrio.

También puede observarse que las personas, que antes trabajaban en ciertos puestos de trabajo, ahora son reemplazadas por equipos tecnológicos que cumplen su función. Ahora se encuentran máquinas que fabrican automóviles o programas que realizan el trabajo de una contestadora telefónica. Este riesgo afecta directamente a la clase trabajadora, aunque puede favorecer económicamente a una empresa.

Hay riesgos en la modernización permanente de los equipos tecnológicos, demandando la actualización, contratación de expertos y demanda de gastos, no previsible.

4. RESULTADOS

La información y su medio de interacción es utilizada como recurso de distribución del conocimiento, la privacidad sufre la arremetida de contenidos que requieren conocer el alcance que tiene dentro de las áreas de negocios, con ello, por medio de redes sociales impulsadas por los servicios prestado por los proveedores e instituciones, se convierten en sitios atractivos para la adquisición de los datos. (Tabla 1).

Tabla 1. Riesgos en el uso de la información en IA

Riesgo	Descripción
Manipulación social	“La información de las personas puede ser utilizada con fines inapropiados” (Carman M., Koerber M., Li J., Choo K. K. R., & Ashman, H., 2018).
Invasión de la privacidad	“Se puede obtener información valiosa de personas o empresas, por medio de proveedores de servicios y redes sociales” (Kanamori S., Nojima, R., Sato H., Tabata N., Kawaguchi K., Suwa H., & Iwai A. 2016).
Contradicción entre objetivo personal y el de la IA	Cuando se desea realizar una acción con un buen fin pero termina perjudicando al que lo aplica.
Discriminación masiva	“Sucedee al aplicar acciones sobre alguien basados en suposiciones, por imágenes o comentarios, registrados en redes sociales o cámaras” (Jiang B., Guo Z., Zhu Q., & Huang G., 2017).

Problemas tecnológicos	Los procesos operativos pueden generar complicaciones que afectan el rendimiento de la IA.
Problemas de interacción	Pueden existir problemas de entendimiento de interacción entre el usuario y la nueva tecnología surgente.
Errores en los modelos	Puede suceder que se usen datos erróneos, de muestras mal obtenidas, que ocasionen resultados no esperados.

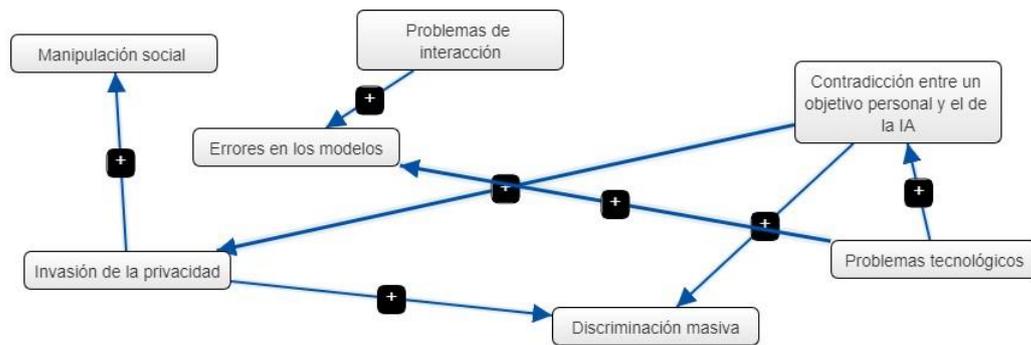


Figura. 3. Diagrama relacionado con los riesgos de la IA.

Se procedió a realizar el siguiente análisis estático y se obtuvo la siguiente tabla de riesgo (Tabla 2).

Tabla 2. Central de factores

	Factor de riesgo	Centralidad	Calificación de riesgo
Manipulación social	0.47	0.47	Receiver
Invasión de la privacidad	0.65	1.69	Ordinary
Errores en los modelos	0.36	1.48	Ordinary
Problemas de interacción	0.47	0.91	Receiver
Discriminación masiva	0.27	0.94	Driver
Concentración entre un objetivo personal y el de la IA	0.48	0.47	Driver
Problemas tecnológicos	0.34	1.2	Receiver

El riesgo de mayor importancia resulta en la invasión de la privacidad con una centralidad de 1.69.

Los nodos son clasificados de la siguiente forma:

Se describe algunos nodos y su impacto en la privacidad, convirtiéndose en uno de los mayores factores dentro del proceso de comunicación (Tabla 3).

Tabla 3. Riesgos en el uso de la información en IA

Riesgo	Tipo
Manipulación social	Receptor
Invasión de la privacidad	Ordinario
Contradicción entre un objetivo personal y el de la IA	Ordinario
Discriminación masiva	Receptor
Problemas tecnológicos	Transmisores
Problemas de interacción	Transmisores
Errores en los modelos	Receptor

5. CONCLUSIÓN

La importancia de utilizar el aprendizaje profundo y redes neuronales útiles y efectivas pueden beneficiar a las empresas, por esto se proyectan a que puedan ser usados por un usuario promedio, pero mientras más se demanda de este tipo de tecnología, mayor cantidad de profesionales especialistas en datos se demanda. Además, hay riesgos conocidos y otros que pueden aparecer de manera inesperada. Se puede trabajar con medidas de seguridad, pero también hay información privada que podría ser encontrada por los hackers y utilizada de manera ilegal.

La IA está comenzando a reemplazando a las personas como mano de obra y eso y eso tiene un gran impacto en la economía de muchas personas. Todos estos aspectos deben de prevenirse para que cause el menor daño posible. El análisis realizado mediante mapas cognitivos difuso permitió priorizar los riesgos plantándose los aspectos relacionados con la privacidad.

REFERENCIAS

- Onik, H. M., Kim, C.-S., & Yang, J. (2019). Personal Data Privacy Challenges of the Fourth Industrial Revolution.
- Leyva-Vázquez, M., Pérez-Teruel, K., Febles-Estrada, A., & Gulín-González, J. (2013). Modelo para el análisis de escenarios basado en mapas cognitivos difusos: estudio de caso en software biomédico. Patel, V. (2017). A Hopfield Neural Network based Reconfiguration Algorithm for Power Distribution Systems.
- Almeida, C. (2019). Análise comparativa de abordagens fuzzy AHP para segmentação de fornecedores sustentáveis com o fuzzy TOPSIS.
- Scott, R. J., Cavana, R. Y., & Cameron, D. (2016). Recent evidence on the effectiveness of group model building. *European Journal of Operational Research*, 249(3), 908–918. <https://doi.org/10.1016/j.ejor.2015.06.078>
- Zhi-Qiang, L. (2001). Causation, Bayesian Networks, and Cognitive Maps.
- Kosko, B., & Burgess, J. (2002). Neural networks and fuzzy systems. *The Mechatronics Handbook*, 3131, 32-1-32–26. <https://doi.org/10.1121/1.423096>
- Vázquez, M. L., & Smarandache, F. (2021). Resolución de sistemas de ecuaciones lineales neutrosóficas mediante computación simbólica. *Revista Asociación Latinoamericana de Ciencias Neutrosóficas*. ISSN 2574-1101, 15, 41–46.
- Smarandache, F., & Leyva-vazquez, M. (2021). UNM Digital Repository THE ENCYCLOPEDIA OF NEUTROSOPHIC RESEARCHERS - 4th Volume (2021).
- Azan Basallo, Y., Estrada Senti, V., & Martinez Sanchez, N. (2018). Artificial intelligence techniques for information security risk assessment. *IEEE Latin America Transactions*, 16(3), 897–901. <https://doi.org/10.1109/TLA.2018.8358671>
- Liu, Y., & Tang, P. (2019). The prospect for the application of the surgical navigation system based on artificial intelligence and augmented reality. *Proceedings - 2018 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2018*, 244–246. <https://doi.org/10.1109/AIVR.2018.00056>
- Carman, M., Koerber, M., Li, J., Choo, K. K. R., & Ashman, H. (2018). Manipulating Visibility of Political and Apolitical Threads on Reddit via Score Boosting. *Proceedings - 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications and 12th IEEE International Conference on Big Data Science and Engineering, Trustcom/BigDataSE 2018*, 184–190. <https://doi.org/10.1109/TrustCom/BigDataSE.2018.00037>
- Kanamori, S., Nojima, R., Sato, H., Tabata, N., Kawaguchi, K., Suwa, H., & Iwai, A. (2016). How Does the Willingness to Provide Private Information Change?
- Jiang, B., Guo, Z., Zhu, Q., & Huang, G. (2019). Dynamic Minimax Probability Machine-Based Approach for Fault Diagnosis Using Pairwise Discriminate Analysis. *IEEE Transactions on Control Systems Technology*, 27(2), 806–813. <https://doi.org/10.1109/TCST.2017.2771732>