



The threads of learning: weaving connections between machines and human minds

Jefferson Rochambrun-Flores, Angel Rivas-Alvarez & Carlos Neyra-Rivera

Facultad de Ingeniería de Sistemas e Informática, Universidad Tecnológica del Perú, Lima, Perú, U20241406@utp.edu.pe, U19207053@utp.edu.pe, c29136@utp.edu.pe

Received: January 24th, 2025. Received in revised form: April 28th, 2025. Accepted: May 12th, 2025

Abstract

Since 2018, OpenAI has led the development of artificial intelligence following the popularization of its ChatGPT tool, promoting research into human-machine communication, which still struggles to represent a person's personality. Therefore, the objective of this study is to identify deep learning methods that can truly recreate human personality. This research used the PICO query, the PRISMA methodology, and a search of the SCOPUS database to identify articles linked to different deep learning approaches to improve human-machine communication. The most notable deep learning models were the mixed model that integrates LSTM with RNN and FNN, the Multi-Distribution Noise algorithm, and Latent Semantic Indexing. It is concluded that deep learning offers powerful tools for human-machine communication but requires continued research to optimize and automate its practical application.

Keywords: machine learning; deep learning; human behavior; artificial intelligence.

Los hilos del aprendizaje: tejiendo conexiones entre máquinas y mentes humanas

Resumen

Desde 2018, OpenAI ha liderado el desarrollo de inteligencia artificial tras la popularización de su herramienta ChatGPT, promoviendo la investigación de la comunicación humano-máquina que sigue teniendo problemas para representar la personalidad de una persona. Por ello, el objetivo del presente estudio es identificar métodos de Deep Learning que puedan recrear genuinamente la personalidad humana. Esta investigación utilizó la pregunta PICO, la metodología PRISMA y la búsqueda en la base de datos SCOPUS para identificar artículos vinculados a diferentes enfoques de aprendizaje profundo para mejorar la comunicación Humano-Máquina. Se identificaron como modelos más destacados de Deep Learning el modelo mixto que integra LSTM con RNN y FNN, el algoritmo Multi-Distribution Noise y el Latent Semantic Indexing. Se concluye que el Deep Learning ofrece herramientas poderosas para la comunicación Humano-Máquina, pero requiere investigaciones continuas para optimizar y automatizar su aplicación práctica.

Palabras clave: aprendizaje mecánico; aprendizaje profundo; comportamiento humano; inteligencia artificial.

1. Introducción

Desde el año 2018, tras el origen de OpenAI (Empresa creadora de ChatGPT) con su modelo de lenguaje GPT-1 se han incrementado los estudios de la comunicación Humano-Máquina (H-M) [1] y esto evidencia el desafío de querer desarrollar una inteligencia artificial (IA) con una personalidad genuina que sea similar al complejo comportamiento humano [2]. El concepto de la comunicación H-M a cobrado

protagonismo con la popularización de la IA y sus capacidades para imitar de forma auténtica a un humano [3], sin embargo, los usuarios aún detectan patrones repetitivos y respuestas artificiales al interactuar con una IA lo que puede generar una experiencia de conversación rígida y restar naturalidad a la interacción [2-4]. Por ello, resulta importante comprender la estrecha relación entre la IA y el Deep Learning (DL), explorar el funcionamiento de las redes neuronales profundas y analizar su capacidad para procesar grandes volúmenes de datos sin la intervención humana [2].

How to cite: Rochambrun-Flores, J., Rivas-Alvarez, A. y Neyra-Rivera, C., Los hilos del aprendizaje: tejiendo conexiones entre máquinas y mentes humanas.. DYNA, 92(237), pp. 130-137, April - June, 2025.



Estos conocimientos son fundamentales para combinar técnicas de IA basadas en distintos enfoques lógicos que potencien las capacidades del DL [3].

A pesar del avance de la IA aún se observa una reducida cantidad de estudios comparativos que analicen de manera rigurosa los métodos que integren el Machine Learning (ML) y DL por lo que aún no se pueden identificar enfoques eficaces para fortalecer la interacción entre H-M [3]. Sin una evaluación profunda de estas técnicas, existe el riesgo de desaprovechar plenamente sus capacidades lo que podría reducir el avance tecnológico por una falta de comprensión en el comportamiento H-M [2,3,5].

La dificultad para lograr una interacción efectiva entre H-M podría deberse a las limitaciones de las redes neuronales profundas ya que a pesar de su eficacia en diversos campos mediante aprendizaje por ensayo y error, pueden fallar catastróficamente al enfrentarse a grandes volúmenes de datos cuya distribución difiere de la utilizada durante el entrenamiento [2,3,5]. El DL (como evolución del ML) utiliza redes neuronales profundas y supera a las redes neuronales estándar mediante la implementación de algoritmos complejos que simulan procesos cognitivos humanos, permitiendo emular la estructura de la lógica humana y derivar a conclusiones a partir del análisis de los datos pero aún presenta limitaciones en su capacidad de adaptación y evolución debido a errores inherentes a su forma de aprendizaje lo que afecta su flexibilidad y podría detener su progreso [3].

La presente investigación busca aportar información para el desarrollo de sistemas más robustos de IA en la comunicación H-M. Diversos artículos han destacado las limitaciones de los modelos tradicionales en el DL y la necesidad de integrar técnicas adicionales para abordar los desafíos presentados por los datos académicos [6]. Hooshyar et al. [5] proponen enriquecer las Redes Neuronales Artificiales (RNA) con conocimiento simbólico con el fin de potenciar su capacidad de generalización y adaptabilidad permitiéndoles aprender de grupos de datos escasos. A pesar del incremento de artículos que abordan estas características, ninguno tiene una visión tan clara sobre el DL [4]. El DL se centraba en la profundidad de las redes neuronales pero este concepto ha evolucionado hasta abarcar técnicas de aprendizaje más complejas y dinámicas [7]. El impacto del DL se ha extendido significativamente impulsando el desarrollo de dominios y aplicaciones que se apoyan de arquitecturas profundas [3]. Por ello, es relevante una investigación que destaque la capacidad de estas técnicas de recopilar datos y en base a ellos aprender sobre el comportamiento humano.

Por lo tanto, el objetivo de la presente investigación es identificar métodos de Deep Learning que puedan recrear genuinamente la personalidad humana, incluso en contextos con un elevado volumen de datos. Es esencial conocer el rendimiento de los modelos ante cambios en las distribuciones de datos, así como también identificar propuestas de soluciones innovadoras que contribuyan a mejorar la comunicación entre H-M.

2. Metodología

Para la presente RSL se realizó una búsqueda de información integral (sin límite de años para las publicaciones) para identificar los métodos de DL que puedan recrear genuinamente la personalidad humana. La búsqueda se realizó empleando la pregunta PIO (Problema, Intervención y Resultados) lo que

permitió realizar una búsqueda más rigurosa en la base de datos seleccionada y como estrategia metodológica se emplearon los lineamientos PRISMA [8].

La pregunta de investigación formulada en la presente investigación fue: ¿Qué patrones de comunicación empleando métodos de Deep Learning se pueden emplear para identificar la comunicación H-M? Con dicha pregunta se generó la siguiente ecuación de búsqueda: ("artificial neural network" OR "deep-learning neural network" OR "unsupervised neural network") AND ((("human behavior" OR "human conduct" OR "human attitude" OR "language technology" OR "technological language") OR ("method comparison" OR "comparison of methods" OR "methods analysis" OR "procedure comparison" OR "comparison of procedures"))). La búsqueda se realizó en la base de datos SCOPUS durante los meses de abril a junio del 2024 empleando los criterios "document type" siendo estos "Article" y "Conference paper". Para la selección de publicaciones se establecieron los siguientes criterios:

Criterios de Inclusión: CI-1: Estudios de neural networks aplicados al aprendizaje, CI-2: Estudios de DL orientados a la comunicación, CI-3: Estudios de métodos o procedimientos aplicados en adultos.

Criterios de Exclusión: CE-1 Estudios de neural networks aplicados a la salud, CE-2 Estudios de Deep Learning orientados a las enfermedades, CE-3 Estudios que se encuentren en un idioma diferente al español e inglés, CE-4: Estudios que no sean de acceso abierto.

Como resultado de la aplicación de la ecuación de búsqueda se identificaron 486 publicaciones las que fueron filtradas conforme al diagrama de flujo PRISMA y los criterios de inclusión y exclusión seleccionándose 10 artículos (Fig. 1).

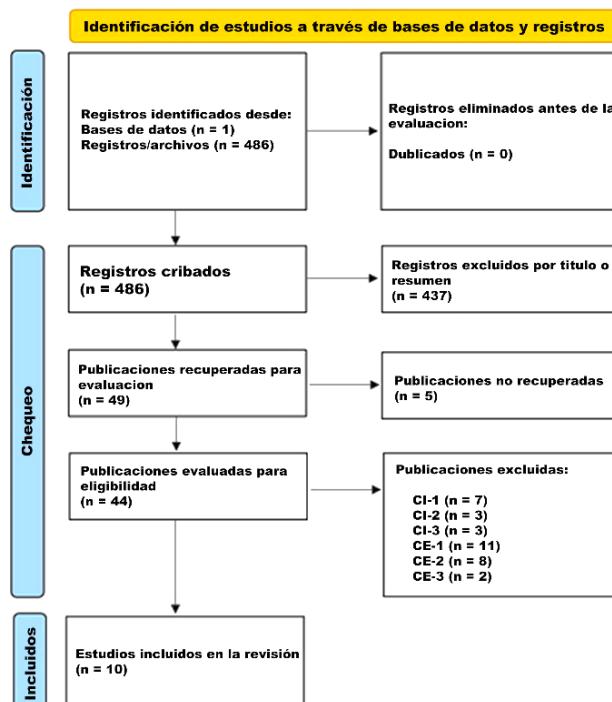


Figura 1. Flujograma PRISMA para la filtración y selección de fuentes.
Fuente: Elaboración propia.

3. Resultados

Los resultados de la presente RSL se han organizado en dos niveles: un análisis descriptivo de los datos bibliométricos y un análisis detallado de las características técnicas relevantes, de acuerdo con el objetivo de esta investigación.

3.1 Análisis bibliométrico

En la Tabla 1 se resumen los principales datos bibliométricos de las investigaciones identificadas. Hasta

Tabla 1.

Datos bibliométricos de artículos seleccionados

Ref/Título/Año	Revista	Citas
[9] Harnessing the flexibility of neural networks to predict dynamic theoretical parameters underlying human choice behavior	2024 PLoS Computational Biology	4
[10] Information architecture applied on natural language processing: a proposal. Information Science contributions on data preprocessing for training and learning of artificial neural networks	2023 Revista Digital de Biblioteconomía e Ciencia da Informação	1
[11] Application of artificial neural networks for personality traits prediction based on handwriting	2023 Indonesian Journal of Electrical Engineering and Computer Science	1
[12] Multi-distribution noise quantisation: an extreme compression scheme for transformer according to parameter distribution	2022 Connection Science	10
[13] Natural language processing with improved deep learning neural networks	2022 Scientific Programming	11
[14] Can we replicate real human behaviour using artificial neural networks?	2022 Mathematical and Computer Modelling of Dynamical Systems	4
[15] On logical inference over brains, behaviour, and artificial neural networks	2021 Computational Brain and Behavior	33
[16] Comparing feedforward and recurrent neural network architectures with human behavior in artificial grammar learning	2020 Scientific Reports	14
[17] Model Detecting Learning Styles with Artificial Neural Network	2019 Journal of Technology and Science Education	45
[18] Experimental research on encoder-decoder architectures with attention for chatbots	2018 Computación y Sistemas	4

Fuente: Elaboración propia

junio del 2024 solo se identificaron 10 publicaciones que cumplían con los criterios de inclusión y exclusión indicados en la presente RSL. El artículo con mayor número de citaciones fue el de Hasibuan, Nugroho y Santosa [17], mientras que las investigaciones con menor cantidad de citas fueron las de Kuroki y Gottschalg-Duque [10] y de Jäger y Reisinger [14].

3.2 Modelos de aprendizaje y sus métodos

El modelo Transformer-based ha demostrado una alta capacidad para alcanzar buenos niveles de precisión en tareas de procesamiento del lenguaje natural siendo empleado por Yu et al. [12], Kuroki y Gottschalg-Duque [10], y Costa-Jussà et al. [18]. Sin embargo, este tipo de modelo requiere elevados recursos computacionales y gran capacidad de almacenamiento debido a la gran cantidad de datos necesarios para su entrenamiento por ello Yu et al. [12] propusieron un enfoque alternativo mediante el método Multi-Distribution Noise (MDN) capaz de generar múltiples resultados posibles junto con sus respectivas probabilidades lo que aligera el procesamiento. Kuroki y Gottschalg-Duque [10] evaluaron el rendimiento de una variante específica (el modelo BERTimbau) analizando su comportamiento al procesar datos con o sin la arquitectura Multimodal Information Architecture (MIA) encontrando que su desempeño varía según el tipo de información utilizada. Costa-Jussà et al. [18] aplicaron un enfoque híbrido entre distintos modelos de redes neuronales desarrollando una arquitectura evolutiva centrada en la codificación y decodificación en el procesamiento del lenguaje natural (NLP) mediante el uso del mecanismo “Attention-based” el que permite al modelo enfocarse en partes específicas de los datos de entrada durante el aprendizaje o la inferencia asignando distintos niveles de importancia a cada segmento para mejorar su rendimiento en tareas como la traducción automática o el procesamiento de imágenes.

El modelo Long Short-Term Memory (LSTM) (orientado al procesamiento del lenguaje natural) fue utilizado por Kuroki y Gottschalg-Duque [10], YiTao [13] y Alamia et al. [16]. Se trata de una variante de Recurrent Neural Networks (RNN), diseñada para capturar y retener dependencias a largo plazo en secuencias. Kuroki y Gottschalg-Duque [10] observaron que el modelo presenta dificultades al captar relaciones entre palabras dentro de una misma oración. Debido a estas limitaciones YiTao [13] propuso un entrenamiento híbrido compuesto por Feed-Forward Neural Network (FNN) para retroalimentación, RNN para modelado secuencial y LSTM como fase final formando así una red evolutiva donde cada componente compensa las debilidades de los demás. Alamia et al. [16] comparan la precisión entre varios modelos cuantificando el entrenamiento de estos y la problemática identificada (incluso con más datos) e identifican que puede ser difícil establecer comparaciones justas con el desempeño humano debido a que los humanos pueden interpretar mejor ciertos tipos de información lo que puede conducir a la generación de personalidades distintas según el modelo y los datos de entrenamiento.

El modelo RNN es un tipo de red neuronal en la que las conexiones entre neuronas forman ciclos lo que permite la

propagación de información en ambas direcciones haciendo que sean adecuadas para manejar datos secuenciales o temporales y conservar estados anteriores. Su uso fue identificado en los trabajos de YiTao [13], Alamia et al. [16], Ger et al. [9] y Costa-jussà [18] quienes lo emplean como una de las fases dentro del entrenamiento híbrido entre el FNN, RNN y LSTM con el objetivo de mejorar la precisión y evitar limitaciones propias de utilizar un solo tipo de red proponiéndose este enfoque como un modelo evolutivo que combina las fortalezas de varias arquitecturas. Alamia et al. [16] observaron que RNN fue el modelo más utilizado en los entrenamientos siendo utilizado en múltiples estudios con diversos conjuntos de datos y estructuras gramaticales. Ger et al. [9] emplearon el entrenamiento de dos variantes del modelo RNN el Recurrent Neural Network Teórica (t-RNN) fundamentada en principios teóricos y matemáticos sólidos cuyo enfoque implicó el estudio detallado de las propiedades internas de las RNN y su aplicación en el diseño de modelos más eficientes y efectivos; y la variante Data-driven Recurrent Neural Network (d-RNN), basada en grandes conjuntos de datos de entrenamiento y en técnicas de aprendizaje automático tanto supervisado como no supervisado que permitieron ajustar los parámetros del modelo para optimizar su rendimiento en tareas específicas. El estudio incluyó una comparativa entre ambas variantes para determinar cuál presentaba un mejor desempeño en función del contexto de aplicación. El estudio incluyó una comparativa entre ambas variantes para evaluar cuál presentaba un mejor desempeño según el contexto de aplicación. Costa-Jussà [18] implementó el modelo RNN dentro de un enfoque híbrido en combinación con el modelo Transformer-based con el objetivo de alcanzar un alto nivel de desempeño en el procesamiento del NPL y simular una personalidad coherente. Para ello se utilizaron procesos de codificación y decodificación aplicados a las frases empleadas durante los entrenamientos lo que permitió modelar patrones de comportamiento y lenguaje más cercanos a los humanos.

El modelo FNN es una red neuronal simple de retroalimentación, capaz de aprender representaciones jerárquicas de datos estructurados y no estructurados a través de múltiples capas. Este modelo fue utilizado por YiTao [13], Alamia et al. [16] y Costa-Jussà et al. [18] dentro de esquemas de aprendizaje híbrido. YiTao [13] lo propuso como una etapa dentro de la evolución de su arquitectura que combinó tres tipos diferentes de redes neuronales. Alamia et al. [16] identificaron que a pesar de realizar 96,000 sesiones de entrenamiento en total un enfoque híbrido que integraba FNN requirió solo 500 entrenamientos mostrando eficiencia en determinadas tareas. Finalmente, Costa-Jussà et al. [18] emplearon un modelo centrado en el NPL en el que el FNN fue sofisticado mediante múltiples capas para potenciar la evolución de su arquitectura principal.

El modelo Artificial Neural Network (ANN) ha sido aplicado en diversos campos del aprendizaje automático siendo utilizado por Guest y Martin [15], Jäger y Reisinger [14], Hasibuan, Nugroho y Santosa [17], y Remaida et al. [11] cada uno con métodos distintos según el objetivo del modelo. Guest y Martin [15] diseñaron su enfoque basándose completamente en el comportamiento cerebral humano con

el propósito de evaluar hasta qué punto una IA puede aprender a replicar dichos patrones extendiendo su aplicación más allá del NPL hacia áreas como el habla y la interpretación de expresiones. Jäger y Reisinger [14] implementaron un modelo Agent-Based Model (ABM) que simula sistemas complejos mediante agentes individuales que interactúan entre sí y con su entorno bajo reglas simples y aunque no se enfoca en NPL este modelo permite que una IA desarrolle una personalidad emergente a partir de dichas interacciones con resultados que varían según el contexto de aplicación. Hasibuan, Nugroho y Santosa [17] propusieron el modelo Latent Semantic Indexing (LSI) que emplea descomposición de valores singulares para detectar significados semánticos latentes en textos lo que permite generar contenidos personalizados a partir de datos fuente y mejorar la experiencia de aprendizaje incluyendo aplicaciones en lenguaje natural. Remaida et al. [11] desarrollaron el modelo Automatic Personality Prediction from Handwriting que es un algoritmo de aprendizaje automático que infiere rasgos de personalidad a partir de características gráficas de la escritura manual, como tamaño o inclinación de las letras y busca que la IA pueda no solo aprender lenguaje natural sino también construir una personalidad propia basada en los datos escritos por individuos.

3.3 Categorías de los algoritmos de aprendizaje

3.3.1 Algoritmos de aprendizaje y procesamiento de datos

Esta categoría agrupa algoritmos cuya principal fortaleza es su capacidad para procesar eficientemente datos complejos y estructurados. Entre ellos se encuentra el modelo FNN propuesto por Alamia et al. [16] el que permite aprender representaciones jerárquicas a través de múltiples capas aplicable tanto a datos estructurados como no estructurados. Hasibuan, Nugroho y Santosa [17] utilizaron el algoritmo LSI destacando por su habilidad para identificar significados y relaciones semánticas ocultas en grandes volúmenes de texto. Remaida et al. [11] emplearon el enfoque Multi-Layer Perceptron que es una variante específica de red neuronal artificial compuesta por múltiples capas de unidades de procesamiento lo que permite modelar patrones complejos mediante operaciones matemáticas. Estos algoritmos mostraron un desempeño sobresaliente en tareas como predicción de personalidad, análisis semántico y clasificación, tal como se refleja en sus niveles de precisión reportados en la Tabla 2.

3.3.2 Algoritmos de modelado y manejo de datos

El algoritmo MDN propuesto por Yu et al. [12] permite modelar datos distribuidos incorporando ruido e incertidumbre probabilística lo que lo hace robusto frente a entradas variables. YiTao [13] implementó el modelo LSTM que es una arquitectura eficaz para procesar secuencias de lenguaje natural que requieren retención de información a largo plazo. Alamia et al. [16] emplearon RNN para tareas de generación y traducción de texto aprovechando su capacidad para capturar dependencias en datos secuenciales. Ger et al.

[9] desarrollaron dos variantes de RNN que fueron t-RNN (basada en fundamentos matemáticos para optimizar la arquitectura) y d-RNN (que se entrena directamente sobre grandes volúmenes para adaptarse a patrones específicos). Ambos enfoques buscan maximizar la eficacia en el modelado secuencial y la representación de comportamientos complejos.

3.3.3 Algoritmos de tratamiento y arquitectura de datos

Kuroki et al. [10] propusieron la arquitectura MIA centrada en el diseño, organización y representación de datos para facilitar su interpretación semántica. Guest y Martin [15] desarrollaron el algoritmo Cognitive Computational Neuroscience (CCN) mediante el cual simularon procesos mentales complejos utilizando modelos computacionales basados en principios neuronales y neurocientíficos. Jäger et al. [14] emplearon el modelo ABM para simular comportamientos emergentes mediante agentes autónomos destacando su utilidad para estudiar la formación de rasgos de personalidad a partir de interacciones sociales. Remaida et al. [11] diseñaron un algoritmo de predicción automática de la personalidad a partir de la escritura manual combinando múltiples fuentes de información gráfica para inferir rasgos psicológicos con técnicas de análisis de datos. Costa-Jussà et al. [18] implementaron el algoritmo Attention-based Mechanisms caracterizado por su capacidad para procesar secuencias extensas, enfocándose en segmentos específicos de entrada y reconociendo estructuras informativas clave que mejoran la comprensión del lenguaje.

3.4 Resultados de entrenamiento y comparativa para la comunicación H-M

Una vez analizados todos los algoritmos, es crucial determinar el grado de precisión que estos algoritmos alcanzaron para recrear la personalidad humana, comenzando por presentar la precisión de algunos de los algoritmos (Tabla 2).

Con esta información se pudo identificar el grado de precisión alcanzado por distintos algoritmos de aprendizaje en la recreación de la personalidad humana, destacando los modelos de Jager et al. [14] y Hasibuan, Nugroho y Santos [17] por su mayor grado de precisión (Fig. 2).

Tabla 2.
Algoritmos de diferentes autores, tareas y precisión promedio.

Ref.	Algoritmos	Tarea	Precisión Promedio
[10]	MIA	Análisis de dominios de conocimiento para generar predicciones	77.02%
[11]	Automatic personality prediction from handwriting	Predecir personalidad en base a la escritura a mano	73.60%
[12]	LSTM-RMC	Cuantificación de datos	31.60%
[12]	RLSTM	Cuantificación de datos	89.46%
[14]	ABM	Agentes que recrean rasgos de la personalidad (Altruismo)	97.00%
[17]	LSI	Aprendizaje de modelos de detección	93.50%

Fuente: Elaboración propia.

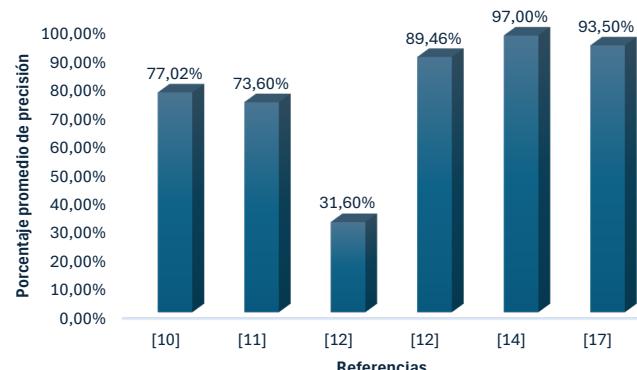


Figura 2: Porcentaje promedio de precisión de los algoritmos identificados
Fuente: Elaboración propia.

Existen algoritmos orientados a mejorar la interacción H-M mediante la recreación más precisa de rasgos de personalidad. Un ejemplo destacado es el algoritmo propuesto por Costa-Jussà et al. [18] quienes demostraron que el uso de mecanismos de atención permite aumentar la precisión del modelo y reducir las limitaciones comunicativas de los chatbots. Esta mejora fue evidenciada a través de evaluaciones externas que calificaron positivamente el desempeño conversacional tras su implementación.

El resultado de este proceso se relaciona con la propuesta de Guest y Martin [15] quienes investigaron el uso ANN en el ámbito de la neurociencia para modelar y comprender procesos cerebrales complejos como la toma de decisiones y el razonamiento metateórico. Su estudio identificó relaciones clave y desafíos en el uso de modelos de aprendizaje profundo para representar funciones cognitivas humanas. Esta perspectiva resulta fundamental para evaluar el rendimiento de los algoritmos y mejorar su capacidad de adaptación en tareas cognitivas avanzadas.

Ger et al [9] demostraron cómo el uso de redes neuronales permite modelar con mayor fidelidad la predicción de acciones y la toma de decisiones humanas, mediante el análisis de datos estructurados a través de sus algoritmos d-RNN y t-RNN.

Las arquitecturas LSTM se destacan por su capacidad para aprender dependencias a largo plazo en secuencias de datos lo que las hace especialmente útiles para capturar patrones complejos del comportamiento humano relacionados con la estructura gramatical, en concordancia con hallazgos previos en percepción visual [8]. En este sentido, Yu et al. [12] y YiTao [13] han resaltado cómo el modelo LSTM contribuye significativamente al entendimiento del lenguaje natural, aunque requiere grandes volúmenes de datos para lograr una representación más fiel de la personalidad humana. Alamia et al. [16] compararon múltiples configuraciones del modelo LSTM y analizaron la cantidad de secuencias de datos necesarias para su entrenamiento (Fig. 3). Sus hallazgos sugieren que, en general, a mayor volumen de datos, mayor es la precisión alcanzada por el modelo.

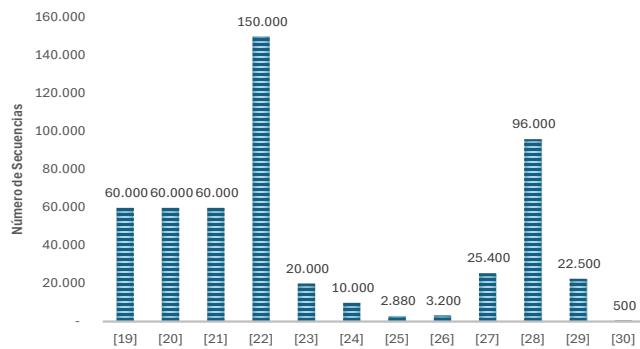


Figura 3: Cantidad de Secuencias de Datos usadas por distintos algoritmos
Fuente: Elaboración propia.

En resumen, los avances en algoritmos de aprendizaje profundo están transformando las interacciones H-M al mejorar significativamente la capacidad de representar con mayor fidelidad la personalidad humana [15]. Estos progresos incluyen tanto modelos basados en mecanismos de atención como redes neuronales recurrentes (como las LSTM) ofreciendo cada uno fortalezas particulares en la captura de comportamientos complejos (Tabla 3) [16]. Además, la cantidad de datos utilizada en el entrenamiento de estos modelos desempeña un rol crucial en su precisión y eficacia (Fig. 3) lo que subraya la necesidad constante de investigación y desarrollo en este campo dinámico y en rápida evolución [9].

Tabla 3.
Comparación de Fortalezas y Debilidades de Modelos de Aprendizaje Profundo

Modelos	Fortalezas	Debilidades
Enfoque Híbrido con LSTM, RNN y FNN	<ul style="list-style-type: none"> Combinación de memoria a largo plazo con el modelado secuencial de RNN y la extracción de características FNN Mejora la coherencia y entiende el contexto en las respuestas Mayor flexibilidad ante distintas clases de datos 	<ul style="list-style-type: none"> Arquitectura compleja que exige mayor poder de cálculo Difícil ajuste de hiperparámetros al integrar varios submodelos Riesgo de sobreajuste si no se cuenta con suficiente volumen de datos
MDN (Multi-Distribution Noise)	<ul style="list-style-type: none"> Permite modelar incertidumbre y variaciones en los datos de entrada Genera múltiples hipótesis de salida con sus probabilidades Robusto frente a ruido y datos heterogéneos 	<ul style="list-style-type: none"> Requiere un diseño cuidadoso de las distribuciones para cada tarea Incrementa la complejidad probabilística del modelo Puede ser inestable si las distribuciones no están bien parametrizadas
LSI (Latent Semantic Indexing)	<ul style="list-style-type: none"> Captura relaciones semánticas latentes más allá de la coincidencia de términos Útil para agrupar textos según subtexto emocional o temático 	<ul style="list-style-type: none"> Pierde eficacia con vocabularios muy amplios o jerga local sin suficiente representación Depende de la calidad y tamaño de la matriz término-documento

Fuente: Elaboración propia

4. Discusión

La presente RSL reveló aspectos fundamentales para el desarrollo de la comunicación H-M mediante técnicas de DL, describiendo como ejemplo métodos para evitar pérdidas en los datos (como Transformer-based) que si bien ofrecen alta precisión exigen una gran cantidad de capacidad computacional, por lo que propusieron enfoques como el método MDN para prevenir gran pérdida de la precisión y menos consumo en almacenamiento y al ser comparado con otros modelos se pudo obtener mejores precisiones con menos datos [12].

También se identificó como una tendencia común en métodos y modelos el uso de modelos híbridos o mixtos que combinan diferentes modelos de red neuronal o métodos dentro de estas bases para compensar las debilidades individuales [12-18]. Este enfoque evolutivo es clave para estos estudios enfocados en la creación de una IA que no solo sea capaz de comunicarse sino que genere vínculos reales con el usuario reproduciendo rasgos de personalidad extraídos de información humana, con la posibilidad de construir incluso una moral emergente basada en dichos datos.

Entre los modelos analizados por Yu et al. [12] se observó que el modelo LSTM-RMC obtuvo un porcentaje de precisión menor al resto de algoritmos (31.60%) utilizando una gran cantidad de conjuntos de datos de WikiText-103 (122900 secuencias de datos) a diferencia de LRLSTM-1500 que empleó PENN Treebank consiguiendo un mayor porcentaje de precisión (89.46%) a pesar de usar una menor cantidad de datos (23,720 secuencias de datos). Uno de los mayores porcentajes de precisión (93.50%) fue obtenido por Hasibuan, Nugroho y Santosa [17] empleando el modelo LSI debido al trabajo en conjunto a las ANN demostrando una mejoría en el nivel de precisión del modelo y su representación de la personalidad humana.

Esto confirma que la calidad de los datos incide directamente en la precisión de los modelos de aprendizaje profundo afectando su capacidad para facilitar una comunicación H-M efectiva y una representación fidedigna de la personalidad humana.

Como se observa en la Fig. 3 la cantidad de datos utilizada por cada modelo puede variar significativamente. Por ejemplo, el modelo de Cohen et al. [29] empleó 225 000 secuencias mientras que el de Alamia et al. [30] trabajó con solo 500. Esta diferencia sugiere que el rendimiento puede depender no solo del volumen, sino también de la calidad del contenido y de la capacidad del algoritmo para extraer patrones significativos.

El análisis comparativo de los tres modelos principales (el enfoque híbrido -LSTM + RNN + FNN-, el algoritmo MDN y la técnica LSI) pone en evidencia tanto sus aportes al campo como las limitaciones y sesgos que pueden surgir en su aplicación a la replicación de la personalidad humana en sistemas de comunicación.

4.1 Complejidad y ajuste de modelos híbridos

Las arquitecturas híbridas combinan la memoria a largo plazo de LSTM, la secuencialidad de RNN y la extracción de características de FNN, logrando respuestas más coherentes

y contextualizadas. Sin embargo, requieren gran poder de cómputo y son complejas de ajustar: cada submodelo introduce hiperparámetros propios, lo que eleva el riesgo de sobreajuste si los datos de habla hispana no abarcan adecuadamente variaciones dialectales y expresiones regionales [13,16].

El MDN modela incertidumbre generando varias hipótesis de salida con sus probabilidades, lo que aumenta la robustez ante datos ruidosos o heterogéneos [12]. No obstante, depende de una parametrización cuidadosa de las distribuciones de lo contrario podría acentuar sesgos mayoritarios del corpus y perder matices culturales de subgrupos menos representados lo que podría limitar su aplicabilidad en contextos culturalmente diversos. El modelo LSI extrae relaciones semánticas latentes mediante SVD [17] pero sólo analiza texto estático sin capturar la dimensión temporal o emocional del discurso, además, en vocabularios extensos o con abundante jerga local se ve afectada la fidelidad de las representaciones semánticas y limita su aplicabilidad a contextos conversacionales reales.

Un reto común a todos los modelos es la representatividad del conjunto de datos. La mayoría de los estudios utilizan corpus en inglés o español formal, dejando fuera expresiones coloquiales y variaciones dialectales regionales lo que introduce sesgos sistemáticos en los resultados [5,6] y limita la capacidad de generalización de los modelos a entornos socioculturales diversos.

Toda esta información demuestra el papel fundamental que desempeñan las redes neuronales en el aprendizaje profundo, al facilitar el procesamiento eficiente de grandes volúmenes de información mediante diversos algoritmos. Asimismo, se destaca la relación planteada por Guest y Martin [17] entre las ANN y la ciencia cognitiva lo que respalda su potencial para simular de manera más realista la personalidad humana.

5. Conclusiones

En conclusión, se identificaron 3 enfoques de aprendizaje profundo con alto potencial para fortalecer la comunicación H-M mediante la recreación genuina de la personalidad, destacando entre ellos el modelo mixto de YiTao [13] que combina LSTM, RNN y FNN en un esquema de retroalimentación para mejorar la precisión y logró resultados más precisos que el resto de modelos al emplear múltiples modelos que se retroalimentan mutuamente para identificar relaciones dentro de la información y recrear mejor la personalidad humana. Hasibuan, Nugroho y Santosa [17] emplearon un modelo basado en LSI que fue capaz de hallar relaciones semánticas en grandes volúmenes de datos y el algoritmo de MDN de Yu et al. [12] se orientó a generar múltiples resultados al comprimir secuencias de datos extensas y reducir el almacenamiento, aumentando la precisión de su análisis mediante un procesamiento más sencillo y sin depender de grandes cantidades de datos. Por esa razón, estos modelos podrían llegar a ser muy buenos al imitar la gran mayoría de aspectos necesarios para hacer una comunicación lo más cercana a una personal real.

Los modelos basados en redes neuronales como Transformer, LSTM y RNN demostraron ser eficaces para

comprender la personalidad humana y la captura de relaciones complejas de datos, permitiendo una mayor precisión debido a su facilidad para la comprensión de información abstracta. Las redes neuronales como ANN mostraron versatilidad al ser aplicadas no solo en NPL sino también se les puede dar un enfoque holístico tanto en la escritura como en el reconocimiento de voz o la reproducción propia de sonido para imitar el habla. La dependencia de grandes conjuntos de datos, la calidad del contenido y la distribución siguen representando un obstáculo para alcanzar este objetivo a pesar de que modelos como LSI han demostrado que es posible lograr buenos resultados sin grandes cantidades de datos, especialmente si se emplean técnicas adecuadas de aprendizaje.

Es esencial continuar investigando y desarrollando técnicas que no solo se enfocuen en la precisión, sino también métodos de aprendizaje mixto que permitan una mayor adaptabilidad en distintos contextos de interacción. Estos hallazgos pueden servir como base para futuras propuestas para optimizar la adaptabilidad de los modelos en diferentes contextos de interacción para mejorar su precisión y la naturalidad de la comunicación entre H-M.

Referencias

- [1] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I., Improving language understanding by generative pre-training. amazon web services, [online]. 2018. Disponible en: <https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/>
- [2] Manakitsa, N., Maraslis, G.S., Moysis, L., and Fragulis, G.F., A review of machine learning and deep learning for object detection, semantic segmentation, and human action recognition in machine and robotic vision. *Technologies*, 12(2), art. 20015, 2024. DOI: <https://doi.org/10.3390/technologies12020015>
- [3] Negro, P., and Pons, C., Artificial Intelligence techniques based on the integration of symbolic logic and deep neural networks: a systematic review of the literature. *Inteligencia Artificial*, 25(69), pp. 13-41, 2022. DOI: <https://doi.org/10.4114/intartif.vol25iss69pp13-41>
- [4] Maroto-Gómez, M., Castro-González, Á., Castillo, J.C., Malfaz, M., and Salichs, M.A., An adaptive decision-making system supported on user preference predictions for human-robot interactive communication. *User Modeling and User-Adapted Interaction*, 33(2), pp. 359–403, 2023. DOI: <https://doi.org/10.1007/s11257-022-09321-2>
- [5] Hooshyar, D., Azevedo, R., and Yang, Y., Augmenting deep Neural Networks with symbolic educational knowledge: towards trustworthy and interpretable AI for education. *Machine Learning and Knowledge Extraction*, 6(1), pp. 593-618, 2024. DOI: <https://doi.org/10.3390/make6010028>
- [6] Wang, P., Fan, E., and Wang, P., Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, 141, pp. 61–67, 2021. DOI: <https://doi.org/10.1016/j.patrec.2020.07.042>
- [7] Yilmaz, A.A., Guzel, M.S., Bostancı, E., and Askerzade, I., A novel action recognition framework based on deep-learning and genetic algorithms. *IEEE Access*, 8, pp. 100631–100644, 2020. DOI: <https://doi.org/10.1109/ACCESS.2020.2997962>
- [8] Page, M.J., McKenzie, J.E., Bossuyt, P.M., et al. Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista panamericana de salud pública*, 46, art. 112, 2020. DOI: <https://doi.org/10.26633/RPSP.2022.112>
- [9] Ger, Y., Nachmani, E., Wolf, L. and Shahar, N., Harnessing the flexibility of neural networks to predict dynamic theoretical parameters underlying human choice behavior. *PLOS Computational Biology*, 20(1), art. 1011678, 2024. DOI: <https://doi.org/10.1371/journal.pcbi.1011678>

- [10] Kuroki-Júnior, G.H., and Gottschalg-Duque, C., Information architecture applied on natural language processing: a proposal. Information Science contributions on data pre-processing for training and learning of artificial neural networks. *Revista Digital de Biblioteconomia e Ciência da Informação*, 21, art. 023002, 2023. DOI: <https://doi.org/10.20396/rdbsci.v21i00.8671396/30919>
- [11] Remaida, A., Abdellaoui, B., Lafraxo, M.A., Sabri, Z., Nouib, H., El Idrissi, Y.E.B., and Moumen, A., Application of artificial neural networks for personality traits prediction based on handwriting. *Indonesian Journal of Electrical Engineering and Computer Science*, 31(3), pp. 1534-1544, 2023. DOI: <http://doi.org/10.11591/ijeecs.v31.i3.pp1534-1544>
- [12] Yu, Z., Li, S., Sun, L., Liu, L. and Haining, W., Multi-distribution noise quantisation: an extreme compression scheme for transformer according to parameter distribution. *Connection Science*, 34(1), pp. 990-1004, 2022. DOI: <https://doi.org/10.1080/09540091.2021.2024510>
- [13] YiTao, Z., Natural language processing with improved deep learning neural networks. *Scientific Programming*, 2022, pp. 1-8, DOI: <https://doi.org/10.1155/2022/6028693>
- [14] Jäger, G., and Reisinger, D., Can we replicate real human behaviour using artificial neural networks?. *Mathematical and Computer Modelling of Dynamical Systems*, 28(1), pp. 95-109, 2022. DOI: <https://doi.org/10.1080/13873954.2022.2039717>
- [15] Guest, O., and Martin, A.E., On logical inference over brains, behaviour, and artificial neural networks. *Computational Brain & Behavior*, 6, pp. 213-227, 2023. DOI: <https://doi.org/10.1007/s42113-022-00166-x>
- [16] Alamia, A., Gauducheau, V., Paisios, D., and VanRullen, R., Comparing feedforward and recurrent neural network architectures with human behavior in artificial grammar learning. *Scientific Reports*, 10(1), pp. 22172, 2020. DOI: <https://doi.org/10.1038/s41598-020-79127-y>
- [17] Hasibuan, M.S., Nugroho, L.E. and Santosa, P.I., Model detecting learning styles with artificial Neural Network. *Journal of Technology and Science Education*, 9(1), pp. 85-95, 2019. DOI: <https://doi.org/10.3926/jots.540>
- [18] Costa-Jussà, M.R., Nuez, Á., and Segura, C., Experimental research on encoder-decoder architectures with attention for chatbots. *Computación y Sistemas*, 22(4), pp. 1233-1239, 2018. DOI: <https://doi.org/10.13053/cys-22-4-3060>
- [19] Cleeremans, A., Servan-Schreiber, D., and McClelland, J.L., Finite state automata and simple recurrent networks. *Neural Computation*, 1, pp. 372-381, 1989. DOI: <https://doi.org/10.1162/neco.1989.1.3.372>
- [20] Servan-Schreiber, D., Cleeremans, A., and McClelland, J.L., Graded state machines: the representation of temporal contingencies in simple recurrent networks. *Machine Learning*, [online]. 7, pp. 161-193, 1991. Available at: <https://link.springer.com/article/10.1007/BF00114843>
- [21] Cleeremans, A., and McClelland, J.L., Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120(3), pp. 235-253, 1991. DOI: <https://doi.org/10.1037/0096-3445.120.3.235>
- [22] Giles, C.L., Miller, C.B., Chen, D., Chen, H.H., Sun, G.Z. and Lee, Y.C., Learning and extracting finite state automata with second-order recurrent neural networks. *Neural Computation*, 4, pp. 393-405, 1992. DOI: <https://doi.org/10.1162/neco.1992.4.3.393>
- [23] Dienes, Z., Connectionist and memory-array models of artificial grammar learning. *Cognitive Science*, 16, pp. 41-79, 1992. DOI: https://doi.org/10.1207/s15516709cog1601_2
- [24] Bodén, M., and Wiles, J., Context-free and context-sensitive dynamics in recurrent neural networks. *Connection Science*, 12(3/4), pp. 197-210, 2000. DOI: <https://doi.org/10.1080/095400900750060122>
- [25] Kinder, A., and Shanks, D.R., Amnesia and the declarative/nondeclarative distinction: a recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience*, 13(5), pp. 648-669, 2001. DOI: <https://doi.org/10.1162/089892901750363217>
- [26] Tunney, R.J., and Shanks, D.R., Subjective measures of awareness and implicit cognition. *Memory & Cognition*, 31, pp. 1060-1071, 2023. DOI: <https://doi.org/10.3758/BF03196127>
- [27] Petersson, K.M., Grenholm, P. and Forkstam, C., Artificial grammar learning and neural networks. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, [online]. 2005, pp. 1726-1731. Available at: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1bfe0e63cefcc9eeffce22c32ae676350b714f6aca>
- [28] Wierzchon, M., and Barbasz, J.A., Wierzchon, M., and Barbasz, J., A connectionist model of artificial grammar learning: simulations based on Higham (1997) Indexes of Knowledge Representation. *29th Annual Conference of the Cognitive Science Society*, 2008.
- [29] Cohen, M., Caciularu, A., Rejwan, I., and Berant, J., Inducing regular grammars using recurrent Neural Networks. *Computation and Language*, pp. 1-5, 2018. DOI: <https://doi.org/10.48550/arXiv.1710.10453>
- [30] Alamia, A., Gauducheau, V., Paisios, D., and VanRullen, R., Which Neural Network Architecture matches Human Behavior in Artificial Grammar” Learning? *Neurons and Cognition*, 2019. DOI: <https://doi.org/10.48550/arXiv.1902.04861>

J. Rochambrun-Flores, es estudiante de último ciclo de la carrera de Ingeniería de Sistemas de la Universidad Tecnológica del Perú.
Orcid: 0009-0003-2195-7469

A. Rivas-Álvarez, es estudiante de último ciclo de la carrera de Ingeniería de Sistemas e Informática de la Universidad Tecnológica del Perú.
ORCID: 0009-0006-7805-7252

C. Neyra-Rivera, Dr. en Biología Molecular y Biotecnología. Investigador en el área de Ciencias de la Salud, Ciencias Básicas y Ciencias Forenses. Ha desarrollado proyectos de investigación a nivel nacional (Perú) e internacional y es docente universitario tanto a nivel de pregrado como de posgrado en asignaturas relacionadas con Biología Celular, Bioquímica e Investigación.
ORCID: 0000-0003-1594-4947