

Diseño de un sistema basado en LLMs para mejorar la toma de decisiones en rehabilitación asistida por robots

Martínez-Pascual, D.^{a,*}, Rubira-Úbeda, P.^a, Reig-Lozano, A.^a, Martín-Batanero, R.^a, Lledó, L. D.^a, García-Aracil, N.^a

^aGrupo de Robótica e Inteligencia Artificial, Instituto de Bioingeniería de Elche, Universidad Miguel Hernández,
Avenida de la Universidad, s/n. 03202 Elche, Alicante, España

Resumen

El uso de dispositivos robóticos de rehabilitación ha surgido como una solución prometedora para mejorar la recuperación motora durante la rehabilitación. Uno de los principales retos en las terapias de rehabilitación es ofrecer una terapia personalizada a cada paciente con el objetivo de maximizar la recuperación motora y cognitiva. Sin embargo, debido a la cantidad de ajustes y datos que ofrecen los dispositivos robóticos, es posible que resulte demasiado complejo para el equipo clínico optimizar los parámetros de la terapia. En este trabajo se describe el diseño de un sistema de apoyo a la toma de decisiones clínicas (CDSS) basado en modelos de lenguaje grandes (LLMs) que permita al personal clínico familiarizarse con los dispositivos, analizar los resultados de las actividades realizadas durante la terapia y realizar sugerencias de los ajustes que deben establecerse para ofrecer una terapia personalizada.

Palabras clave: Tecnología de asistencia e ingeniería de rehabilitación, Ingeniería de rehabilitación y asistencia sanitaria, Toma de decisiones y procesos cognitivos, Automatización centrada en el ser humano, Interfaces inteligentes

Design of an LLM-based system to improve decision-making in robot-assisted rehabilitation.

Abstract

The use of robotic rehabilitation devices has emerged as a promising solution to improve motor recovery during rehabilitation. One of the main challenges in rehabilitation therapies is to provide personalized therapy to each patient to maximize motor and cognitive recovery. However, due to the amount of settings and data offered by robotic devices, it may be too complex for the clinical team to optimize therapy parameters. This paper describes the design of a clinical decision support system (CDSS) based on large language models (LLMs) that allows the clinical staff to familiarize themselves with the devices, analyze the results of the activities performed during therapy, and make suggestions for the settings that should be set to provide personalized therapy.

Keywords: Assistive technology and rehabilitation engineering, Rehabilitation engineering and healthcare delivery, Decision making and cognitive processes, Human centred automation, Intelligent interfaces

1. Introducción

De acuerdo con la Organización de las Naciones Unidas (ONU) se prevé que la proporción de personas de 65 años o más aumente en todo el mundo para 2050 (Desa, 2019). Este segmento de la población es especialmente propenso a sufrir un accidente cerebrovascular (ACV) o ictus, ya que la incidencia relativa de ictus se duplica cada década a partir de los 55 años (Norrving et al., 2018). A pesar de los continuos avances en el tratamiento de los ACV y el desarrollo de nuevas estrategias de rehabilitación, el ictus sigue siendo un acontecimiento

con un gran impacto en la salud pública, que causa 5 millones de muertes al año y un número mucho mayor de supervivientes que viven con discapacidad física y cognitiva crónica. Cabe destacar que de aquellos pacientes que presentan hemiparesia de miembro superior en el momento del ingreso hospitalario, sólo el 8-20 % recupera la función completa tras el tratamiento, y el 50 % tiene problemas con la función del brazo cuatro años después del ictus (Kwakkel et al., 2003). Estos resultados explican claramente por qué la búsqueda de nuevas estrategias para mejorar la recuperación funcional tras un ictus ha sido y sigue

*Autor para correspondencia: david.martinezp@umh.es
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

siendo un elemento central de la investigación en rehabilitación (Langhorne et al., 2009; Wolf et al., 2021). Con el fin de aumentar la eficacia y acelerar el proceso de rehabilitación, esta brecha se cubre a menudo mediante ejercicios asistidos que se orientan hacia el reaprendizaje de movimientos similares a los utilizados en el contexto de la vida diaria. En este sentido, los robots de rehabilitación han demostrado ser una herramienta eficaz para este tipo de ejercicios asistidos.

Asimismo, una característica clínica de los pacientes afectados por el Ictus es un cuadro clínico y una recuperación motora muy variables (Salvalaggio et al., 2024; Favetta et al., 2023). De hecho, debido a la heterogeneidad del cuadro clínico y a la variedad de variables implicadas en el proceso de recuperación, la respuesta a una intervención de rehabilitación puede variar enormemente entre pacientes. Esto sugiere que no existe un enfoque único para todos los pacientes.

Los robots de rehabilitación pueden emplearse para proporcionar terapias personalizadas que cumplan los principios fundamentales de la neurorrehabilitación, es decir, repetición, alta intensidad y especificidad de la tarea. De acuerdo con los principios mencionados, la rehabilitación robótica pretende promover la neuroplasticidad y, por tanto, la recuperación neurofisiológica, induciendo la plasticidad funcional o estructural en las redes cerebrales, controlando tanto las funciones motoras como las cognitivas (Turner et al., 2013). Por otra parte, los dispositivos robóticos también pueden actuar como herramientas de evaluación para cuantificar objetivamente el rendimiento motor de los pacientes (Nordin et al., 2014; Catalán et al., 2023). Gracias a los sensores y actuadores incorporados se puede adquirir datos sobre la cinemática y la cinética de la extremidad superior de los pacientes, que se procesan para obtener índices cuantitativos relacionados con la calidad del movimiento de la extremidad superior. No obstante, la gran cantidad de datos recogidos por estos dispositivos, junto con una posible falta de familiarización del personal clínico con estos, puede hacer que la evaluación del paciente y la elección de las configuraciones de los dispositivos resulte demasiado compleja para que se ajuste de forma óptima los parámetros de la terapia (Labruyère, 2022). Por ello, los sistemas de Inteligencia Artificial (IA) podrían ayudar a gestionar esa complejidad integrando adecuadamente los datos adquiridos junto con protocolos clínicos para ofrecer terapias asistidas por robots personalizadas.

En este trabajo se describe el primer diseño e implementación de un sistema de apoyo a la toma de decisiones clínicas (CDSS, *Clinician Decision Support System*) basado en modelos de lenguaje grandes (LLMs, *Large Language Models*). El objetivo del CDSS propuesto es ayudar al personal clínico familiarizarse con los dispositivos robóticos, analizar los resultados de las actividades realizadas durante la terapia y realizar sugerencias de los ajustes que deben establecerse en los dispositivos robóticos para ofrecer una terapia personalizada.

2. Aplicación de LLMs en el ámbito de la rehabilitación

Uno de los grandes retos en la planificación de los tratamientos de rehabilitación es integrar las teorías generales del aprendizaje motor y los resultados de la investigación sobre el ictus junto con los protocolos de los centros clínicos, el criterio de médicos y terapeutas, el diagnóstico de cada paciente

y la evolución en el tratamiento. Por ello, el desarrollo de un CDSS exige un análisis exhaustivo y una integración eficaz de una gran cantidad de información en formato de texto.

Los LLMs son modelos de IA diseñados para procesar y generar texto en lenguaje natural (Zhao et al., 2024), los cuales pueden tener un gran potencial como motor cognitivo del CDSS propuesto para terapias de rehabilitación. Cabe destacar que el entrenamiento de este tipo de modelos es un proceso altamente costoso, tanto en términos computacionales como económicos. La inversión en hardware especializado, consumo energético y costes de infraestructura ha aumentado significativamente en los últimos años, con estimaciones que indican que entrenar modelos de vanguardia como GPT-4, Gemini o LLaMA puede costar cientos de millones de dólares (Cottier et al., 2025). Sin embargo, los modelos abiertos preentrenados por grandes fundaciones ofrecen una alternativa viable, permitiendo a investigadores y desarrolladores aprovechar estas arquitecturas para tareas específicas sin incurrir en los costos exorbitantes de entrenamiento desde cero (Sharir et al., 2020). Esta estrategia no solo reduce la inversión necesaria, sino que también democratiza el acceso a tecnologías avanzadas de procesamiento del lenguaje natural, facilitando su aplicación en diversos ámbitos.

A pesar de que los LLMs exhiben un notable nivel de conocimiento en diversos dominios gracias a sus enormes conjuntos de datos de preentrenamiento, estos conocimientos tienen dos limitaciones importantes. En primer lugar, el conocimiento es estático y no se actualiza con el tiempo. En segundo lugar, no es específico y, por lo tanto, puede carecer de conocimientos matizados en dominios concretos. Para poder aumentar o adaptar el conocimiento de los LLMs preentrenados existen dos técnicas principales: fine-tuning y el aprendizaje en contexto (Ovadia et al., 2024; Balaguer et al., 2024). Con el fine-tuning se continúa el proceso de aprendizaje del modelo usando datos sobre una tarea específica, con el objetivo de que los pesos del modelo se adapten al conocimiento expuesto, y suele ser altamente efectivo. No obstante, el coste inicial de ajustar un modelo mediante fine-tuning es elevado debido al trabajo que supone elaborar un dataset apropiado con datos nuevos, que para el caso que nos atañe sería en el campo de la rehabilitación. Por otro lado, la base del aprendizaje en contexto consiste en mejorar el rendimiento de los LLMs en nuevas tareas modificando la consulta del usuario al introducirla al modelo, sin ajustar sus pesos. Dentro de esta técnica se encuentra la Generación Aumentada por Recuperación (RAG, *Retrieval Augmented Generation*), cuyo funcionamiento consiste en utilizar una base de conocimiento auxiliar para encontrar documentos similares a la consulta de entrada (Lewis et al., 2020). Para ello, se emplea un modelo de embeddings preentrenado, que transforma la base de conocimiento auxiliar en representaciones matemáticas (vectores) que capturan el significado y las relaciones semánticas entre palabras, frases y documentos. Al convertir tanto la consulta del usuario como los documentos almacenados en vectores, el sistema puede comparar estas representaciones y determinar qué documentos tienen la mayor similitud semántica con la consulta. Este proceso permite recuperar la información más relevante y enriquecer el contexto que se le proporciona al modelo de lenguaje. La principal ventaja de la RAG reside en el bajo coste inicial de crear embeddings para aumentar el conocimiento y precisión de los LLMs. Como punto negativo de esta técnica

ca, cabe mencionar que el tamaño de tokens introducidos como entrada puede incrementar el tamaño del prompt, lo que puede aumentar consigo el coste de las consultas realizadas.

Si bien es cierto que existen diferentes propuestas de uso de LLMs en el ámbito médico, F. Liu et al. destacan que los desarrollos actuales se centran en medicina general, lo que resulta en una baja representación de campos especializados como la terapia de rehabilitación (Liu et al., 2025). Por tanto, resulta fundamental impulsar la investigación y el desarrollo de modelos de lenguaje que se adapten específicamente a las necesidades de la terapia de rehabilitación. La incorporación de estos modelos en la práctica clínica permitiría integrar de manera eficiente el conocimiento especializado, optimizando la toma de decisiones y personalizando las intervenciones según las características individuales de cada paciente. Además, la integración de dispositivos y tecnologías avanzadas, como sistemas robóticos de rehabilitación, sensores biomecánicos y plataformas de telemedicina, representa un paso crucial para mejorar la eficacia de los tratamientos.

Por tanto, el uso de LLMs en este contexto podría facilitar la interpretación de los datos recopilados por dispositivos de rehabilitación, generando recomendaciones más precisas y ajustadas a la evolución de los pacientes. Asimismo, el acceso a información relevante permitiría al personal clínico tomar decisiones fundamentadas, reduciendo la incertidumbre en el diseño de los tratamientos y maximizando el impacto de las terapias.

3. Materiales y Métodos

3.1. Ejecución de LLMs y modelos de embeddings

Uno de los objetivos de este trabajo es proponer un CDSS eficiente, basada en modelos de código abierto y que puedan ejecutarse en equipos o servidores locales para garantizar la privacidad de los pacientes. Para la ejecución del LLM y modelo de embeddings seleccionados se ha empleado el software LMStudio. Este software permite ejecutar de forma sencilla modelos de código abierto empleando el backend llama.cpp, el cual optimiza la inferencia de LLMs en CPU y GPU locales. Además, una de las ventajas de LMStudio es que ofrece una API compatible con el endpoint de OpenAI, lo que la hace fácilmente integrable con diferentes frameworks. En concreto, la implementación del CDSS se ha desarrollado con el framework LlamaIndex (Liu, 2022) en Python.

3.2. Construcción de base de datos de conocimiento sobre un dispositivo robótico de rehabilitación

Como se ha introducido, uno de los principales propósitos del CDSS que se pretende desarrollar es permitir al personal clínico familiarizarse con los dispositivos de los que disponen, además de facilitar el análisis de resultados y sugerencias de ajustes basados en el historial del usuario para optimizar la terapia. Con el fin de plantar las bases del CDSS y analizar su viabilidad, se ha decidido emplear información acerca del robot de rehabilitación Rubidium (Figura 1), el cual ha sido desarrollado por el grupo de investigación y empleado en diversos estudios previos del grupo (Catalán et al., 2023; Martínez-Pascual et al., 2025).



Figura 1: Imagen de un usuario empleando el robot Rubidium durante una terapia de neurorehabilitación. En la Figura también se han representado algunos de los juegos incluidos con el dispositivo.

Este sistema robótico está formado por un paralelogramo articulado accionado por dos motores eléctricos. El sistema permite realizar movimientos en un plano horizontal, y consta de una órtesis anclada al efector final que permite soportar el brazo del usuario. El sistema cuenta con diferentes juegos serios que permiten trabajar la rehabilitación tanto motora como cognitiva. Además, incluye varios modos de asistencia para que pueda ser empleado por pacientes con diferentes grados de afectación motora.

Para la creación de la base de conocimiento acerca del dispositivo, se ha incluido toda la información relevante acerca del robot Rubidium, especificando los modos de asistencia e información acerca de los juegos incluidos, describiéndolos y especificando su tipología: juegos punto a punto, de detección de estímulos visuales y atención, de precisión y control del movimiento, cognitivos, o de movimiento libre y continuo. Asimismo, en la base de conocimiento se incluyen protocolos de uso del robot Rubidium redactados por una terapeuta con amplia experiencia en el uso del dispositivo. Dichos protocolos establecen guías de uso del dispositivo en base a una valoración del paciente mediante la escala clínica de Fugl-Meyer de miembro superior y a una valoración obtenida mediante el propio dispositivo, así como unas directrices a seguir según la evolución y tolerancia del paciente al tratamiento. Entre estas guías se encuentran cómo establecer el rango de movimiento del paciente, qué tipo de modos de asistencia son más adecuados, y qué tipología de juegos aplicar según la evolución y preferencias del paciente (variedad, intensidad, etc.).

Una vez recopilada la información técnica acerca del Rubidium y los protocolos de uso, se empleó el modelo de embeddings para almacenar esta información como vectores en una base de datos de ChromaDB (ver Figura 2.(a)).

3.3. Arquitectura del sistema de apoyo de toma de decisiones

La arquitectura del CDSS está formada por tres flujos de trabajo diferentes implementados en LlamaIndex, los cuales se han representado en la Figura 2 y se describen en esta sección.

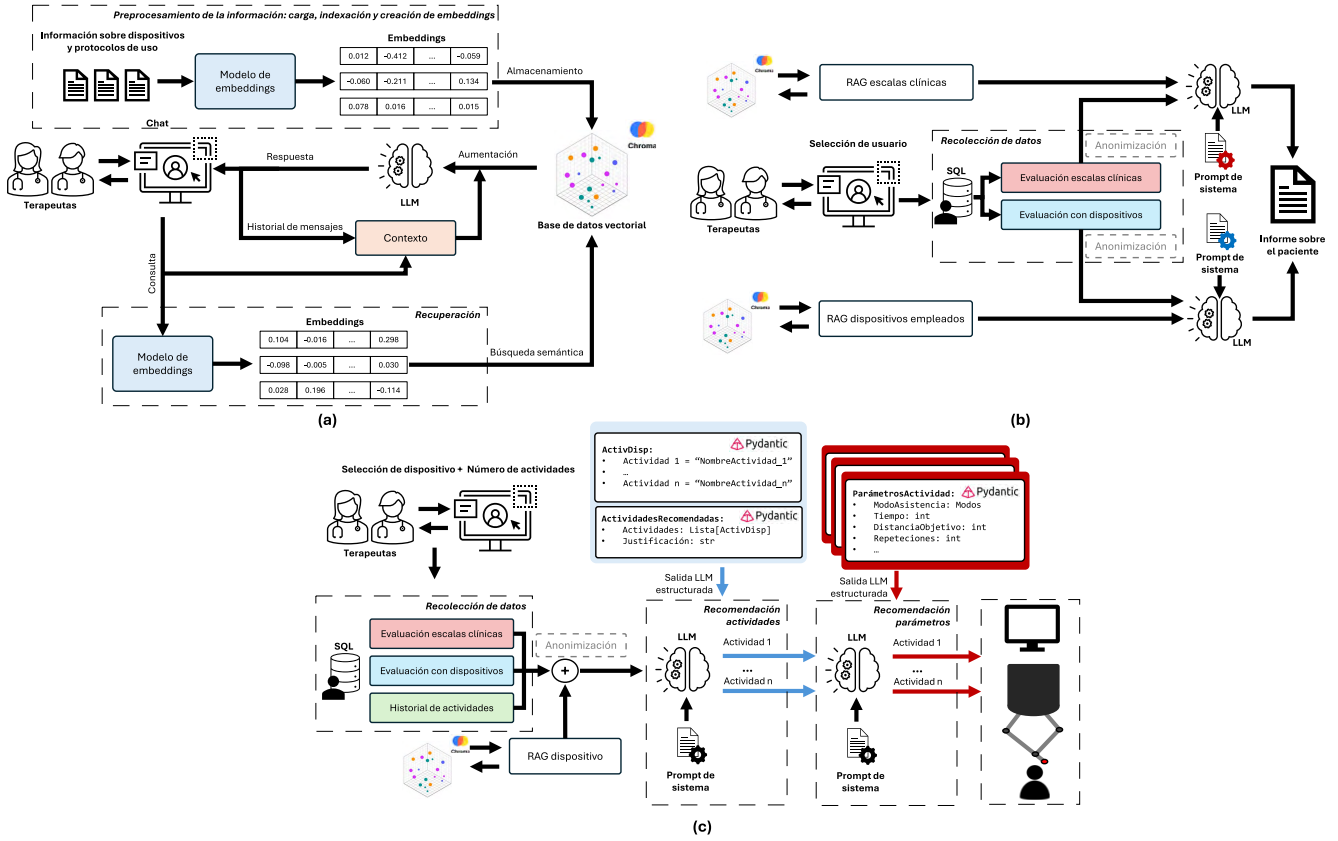


Figura 2: Representación de los diferentes flujos de trabajo implementados en el prototipo de CDSS. (a) Flujo de trabajo del chat que permite interactuar con el LLM para preguntas y respuestas acerca de las características de los dispositivos y protocolos a seguir. (b) Flujo de trabajo para el análisis del historial de pacientes de evaluación de escalas clínicas y resultados de actividades con los dispositivos. (c) Flujo de trabajo para recomendar actividades y sus ajustes.

3.3.1. Chat para familiarización con dispositivos

El flujo de trabajo del sistema de chat para terapeutas con acceso a un LLM y RAG se compone de varias fases fundamentales, el cual se representa en la Figura 2.(a). En primer lugar, se lleva a cabo la indexación y almacenamiento de información, donde los datos técnicos y clínicos del robot de rehabilitación Rubidium se han transformado en representaciones vectoriales mediante el modelo de generación de embeddings Granite-Embedding-278m-multilingual. Estos vectores, que encapsulan el conocimiento sobre los modos de asistencia, actividades y protocolos de uso del dispositivo, se almacenan en una base de datos vectorial, lo que permite una recuperación eficiente y contextualizada de información relevante para los terapeutas.

Cuando un terapeuta inicia una consulta a través del chat, su pregunta es procesada y convertida en embeddings mediante el mismo modelo de generación de representaciones vectoriales. Posteriormente, estos embeddings son utilizados para realizar una búsqueda semántica en la base de datos, identificando los fragmentos de información más pertinentes para la consulta específica. La base de datos vectorial devuelve los resultados más relevantes al LLM, que analiza los datos y genera una respuesta clara y fundamentada, teniendo en cuenta el contexto de la conversación y la información almacenada.

3.3.2. Sistema de apoyo para el análisis de datos

La segunda funcionalidad incluida en el CDSS se ha desarrollado como una ayuda para el análisis de las actividades realizadas así como la evolución del paciente. El proceso que

da soporte a esta funcionalidad se representa en la Figura 2.(b), el cual consta de varios pasos. En la primera parte del flujo de trabajo, los terapeutas seleccionan el usuario de la interfaz del sistema y los datos correspondientes se recuperan de una base de datos estructurada SQL. Los datos recuperados incluyen tanto la valoración con escalas clínicas como los resultados de evaluación con dispositivos. En el caso del robot Rubidium, se incluye toda la información referente a los parámetros empleados durante cada actividad (e.g. modo de asistencia empleado durante las actividades, número de repeticiones, etc.) así como los resultados de las actividades (e.g. porcentaje de acierto).

Antes de procesar la información, se lleva a cabo un proceso de anonimización para garantizar la privacidad de los datos del paciente. Posteriormente, los datos se analizan mediante dos procesos diferentes y simultáneos. En ambos procesos, se recupera información mediante RAG sobre las escalas clínicas empleadas para realizar la evaluación del paciente, así como información relevante del dispositivo Rubidium. Además, la llamada al LLM para analizar los datos incluye prompts del sistema diferentes, adaptados a la información que se va a analizar, pero la instrucción dada es que se debe analizar de forma objetiva la evolución de los resultados, sin incluir sugerencias. Además, se especifica que el informe se genere en formato markdown y situado entre los tokens `< /report >` que indiquen el inicio y final de los respectivos informes.

Finalmente, se postprocesa la salida del LLM y como resultado se junta la información para generar un informe clínico que pueda ser utilizado por terapeutas como un apoyo para

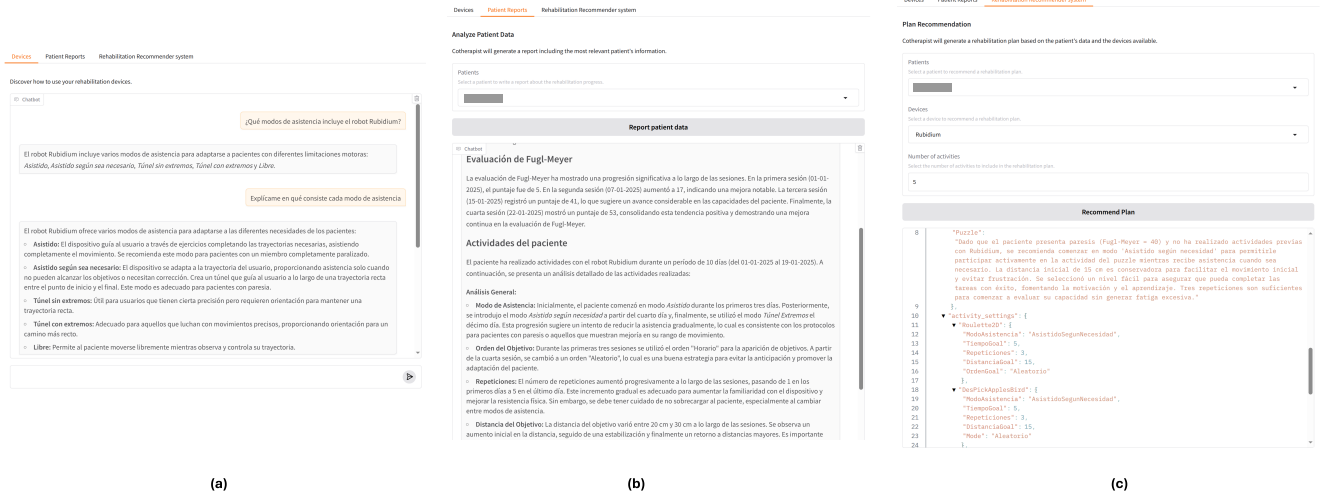


Figura 3: Ejemplos de ejecución del prototipo de CDSS implementado. (a) Chat para familiarización con dispositivos (b) Sistema de apoyo para el análisis de datos (c) Sistema de recomendación de ajustes de los dispositivos

comprender mejor el estado del paciente.

3.3.3. Sistema de recomendación de ajustes de los dispositivos

El proceso seguido para implementar el sistema de recomendación de actividades y sus ajustes se ha representado en la Figura 2.(c). En primer lugar, se debe seleccionar el paciente para quién se realizará la recomendación, la cantidad de actividades de la sesión y el dispositivo que se desea emplear (e.g. robot Rubidium). Tras esta selección, se hace una consulta a la base de datos SQL y se recuperan los datos del paciente. Además, mediante RAG se recupera información sobre los protocolos de uso del dispositivo seleccionado.

Con la información recogida sobre el paciente y el dispositivo se realiza una recomendación de actividades y sus ajustes en dos pasos. Cabe destacar que existe una diferencia principal con respecto a las otras funcionalidades implementadas, y radica en la forma en la que se estructura la salida del LLM.

Tanto en el chat como el sistema de apoyo de análisis de datos tienen por objetivo ofrecer información sobre el dispositivo Rubidium y los datos recogidos. Por tanto, la información se ofrece de manera no estructurada, en forma de lenguaje natural. Sin embargo, para poder realizar una recomendación y establecer los parámetros de las actividades es necesario que la salida del LLM se encuentre estructurada. Para ello, se hace uso de la librería Pydantic, que permite definir modelos de datos precisos y validar la información generada por el modelo de lenguaje. Mediante Pydantic, los datos generados se organizan en estructuras bien definidas, asegurando su coherencia. Esto garantiza que los parámetros sugeridos (e.g. la cantidad de repeticiones, el modo de asistencia, la duración de la actividad, etc.), sean representados de manera consistente y puedan ser aplicados directamente en el dispositivo Rubidium. Gracias a esta estructuración, se mejora la interoperabilidad entre los distintos módulos del sistema.

Como cada actividad del robot Rubidium tiene unos parámetros específicos, es necesario realizar en primer lugar una recomendación de las n actividades que se van a realizar durante la sesión. Tras la llamada del LLM donde se incluye

la información del paciente y el dispositivo, se genera una salida estructurada con dos campos. El primer campo es una lista con las n actividades recomendadas, y el segundo campo es una justificación de las actividades seleccionadas que puedan servir como realimentación para el personal clínico. Siguiendo el mismo procedimiento, para cada una de las actividades recomendadas se genera una salida estructurada con la recomendación de los parámetros de la actividad junto con una justificación de la selección en forma de lenguaje natural.

4. Resultados y discusión

La ejecución del prototipo del CDSS desarrollado se ha llevado a cabo en un equipo local con procesador Intel Core i7-13700K, 32 GB de RAM y GPU Nvidia RTX 4070 Ti 12 GB GDDR6X, empleando el modelo Gemma 3-12B-Instruct-QAT con cuantificación $Q4_0$ y el modelo de embeddings Granite-Embedding-278m-multilingual. La ejecución del LLM genera 17.78 tokens/s con un retraso de 0.21s hasta la generación del primer token, lo que puede considerarse como unas medidas de tiempo de ejecución asumibles.

En la Figura 3 se muestran ejemplos de ejecución del CDSS empleando una interfaz web desarrollada con Gradio. Una de las principales ventajas de Gradio es la facilidad de desarrollo y despliegue. Además, la interfaz incluye una API que permite interactuar con los elementos de la interfaz, lo que facilitaría la integración con una plataforma software de rehabilitación con diferentes dispositivos.

En la Figura 3.(a) se observa como mediante una interfaz de chat el CDSS permite realizar consultas para conocer información sobre el robot de rehabilitación Rubidium. Por otra parte, en la Figura 3.(b) se muestra un ejemplo del informe generado sobre los datos almacenados de un paciente, dónde se describe la evolución de la valoración en la escala Fugl-Meyer de miembro superior y sobre las actividades realizadas con el dispositivo Rubidium. Por último, la Figura 3.(c) muestra la salida en formato JSON del sistema de recomendación de actividades, donde se observa la justificación de elección de los parámetros del juego y la salida estructurada. De esta forma, la estructuración

de los parámetros en formato JSON podrá servir para establecer los parámetros de la actividad de forma sencilla a través de una llamada a la API de la interfaz de Gradio.

Finalmente, cabe destacar que una de las principales limitaciones del CDSS presentado es que por ahora no ha sido evaluado en términos de fidelidad ni revisado por terapeutas expertos en rehabilitación. La plataforma se encuentra en una fase inicial de desarrollo, y si bien integra procesos clave basados en LLMs para el apoyo en la toma de decisiones en la rehabilitación asistida por robots, su eficacia y veracidad en contextos clínicos reales aún deben ser validadas. En futuros trabajos, se evaluará el rendimiento de varios LLMs y modelos de embeddings con el objetivo de optimizar la metodología RAG. Para ello, se emplearán varias métricas (e.g. fidelidad, relevancia, precision y recall del contexto, etc.) y también se valorará implementar una metodología de RAG basada en grafos, las cuales han demostrado tener una gran precisión (Procko and Ochoa, 2024). Asimismo, se llevará a cabo una evaluación rigurosa de la precisión de sus recomendaciones de las actividades y parámetros recomendados, comparando su rendimiento con otros sistemas de recomendación basado en otro tipo de modelos de Machine Learning.

5. Conclusiones

En este trabajo se ha planteado la arquitectura de un novedoso CDSS basado en LLMs que permita facilitar al personal clínico el uso de nuevos sistemas de rehabilitación, como pueden ser las tecnologías robóticas. Para aumentar el conocimiento del LLM empleado, se ha empleado la metodología RAG, creando una base de datos vectorial donde se ha almacenado las guías y protocolos de uso del robot de rehabilitación Rubidium. Gracias a ello, se ha podido crear un chat que permita al personal clínico realizar preguntas al LLM para conocer las características del dispositivo de forma sencilla. Asimismo, haciendo uso de estas tecnologías y accediendo a los datos de los pacientes almacenados en una base de datos, se ha desarrollado una funcionalidad que facilite la creación de informes y el análisis de los datos recogidos por los dispositivos de rehabilitación. Por último, se ha propuesto un sistema de recomendación para realizar sugerencias de qué actividades realizar y qué parámetros establecer (e.g. modo de asistencia, repeticiones, etc). Este sistema de recomendación integra el conocimiento del protocolo de uso del robot Rubidium junto con los datos de actividades y evaluación del paciente para realizar.

El CDSS propuesto todavía se encuentra en fase de desarrollo, y hasta la fecha no ha podido ser validado por terapeutas expertos en la materia para evaluar la fidelidad del sistema. No obstante, consideramos que este primer prototipo puede representar un avance significativo en el uso de tecnologías de IA a la vanguardia en el campo de la rehabilitación.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Universidades a través de la beca de Formación de Profesorado Universitario FPU20/05137, por la Conselleria de Educación, Cultura, Universidades y Empleo de la Comunidad Valenciana a través de la ayuda CIPROM/2022/12 y por la Agencia Estatal de Investigación a través de los proyectos TED2021-130431B-I00 y PID2022-139957OB-I00.

Referencias

- Balaguer, A., Benara, V., Cunha, R. L. d. F., Hendry, T., Holstein, D., Marsman, J., Mecklenburg, N., Malvar, S., Nunes, L. O., Padilha, R., et al., 2024. Rag vs fine-tuning: Pipelines, tradeoffs, and a case study on agriculture. *arXiv preprint arXiv:2401.08406*.
- Catalán, J. M., Blanco-Ivorra, A., García-Pérez, J. V., Vales, Y., Martínez-Pascual, D., Ezquerro, S., Garrote, A., Costa, T., Lledó, L. D., García-Aracil, N., 2023. Patients' physiological reactions to competitive rehabilitation therapies assisted by robotic devices. *Journal of NeuroEngineering and Rehabilitation* 20 (1), 41.
- Cottier, B., Rahman, R., Fattorini, L., Maslej, N., Besiroglu, T., Owen, D., 2025. The rising costs of training frontier ai models. URL: <https://arxiv.org/abs/2405.21015>
- Desa, U., 2019. United nations department of economic and social affairs. Population Division. World Population Prospects.
- Favetta, M., Romano, A., Valè, N., Cieslik, B., Federico, S., Girolami, A., Mazzarotto, D., Pregolato, G., Righetti, A., Salvalaggio, S., et al., 2023. A scoping review of scientific concepts concerning motor recovery after stroke as employed in clinical trials. *Frontiers in Neurology* 14, 1221656.
- Kwakkel, G., Kollen, B. J., van der Grond, J., Prevo, A. J., 2003. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. *Stroke* 34 (9), 2181–2186.
- Labruyère, R., 2022. Robot-assisted gait training: more randomized controlled trials are needed! or maybe not? *Journal of neuroengineering and rehabilitation* 19 (1), 58.
- Langhorne, P., Coupar, F., Pollock, A., 2009. Motor recovery after stroke: a systematic review. *The Lancet Neurology* 8 (8), 741–754.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-t., Rocktäschel, T., et al., 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems* 33, 9459–9474.
- Liu, F., Zhou, H., Gu, B., Zou, X., Huang, J., Wu, J., Li, Y., Chen, S. S., Hua, Y., Zhou, P., et al., 2025. Application of large language models in medicine. *Nature Reviews Bioengineering*, 1–20.
- Liu, J., 11 2022. LlamaIndex. URL: https://github.com/jerryjliu/llama_index DOI: 10.5281/zenodo.1234
- Martínez-Pascual, D., Catalán, J. M., Lledó, L. D., Blanco-Ivorra, A., García-Aracil, N., 2025. A deep learning model for assistive decision-making during robot-aided rehabilitation therapies based on therapists' demonstrations. *Journal of NeuroEngineering and Rehabilitation* 22 (1), 18.
- Nordin, N., Xie, S. Q., Wünsche, B., 2014. Assessment of movement quality in robot-assisted upper limb rehabilitation after stroke: a review. *Journal of neuroengineering and rehabilitation* 11, 1–23.
- Norring, B., Barrick, J., Davalos, A., Dichgans, M., Cordonnier, C., Guekht, A., Kutluk, K., Mikulik, R., Wardlaw, J., Richard, E., et al., 2018. Action plan for stroke in europe 2018–2030. *European stroke journal* 3 (4), 309–336.
- Ovadia, O., Brief, M., Mishaali, M., Elisha, O., 2024. Fine-tuning or retrieval? comparing knowledge injection in llms. URL: <https://arxiv.org/abs/2312.05934>
- Procko, T. T., Ochoa, O., 2024. Graph retrieval-augmented generation for large language models: A survey. In: 2024 Conference on AI, Science, Engineering, and Technology (AIxSET). IEEE, pp. 166–169.
- Salvalaggio, S., Gambazza, S., Andò, M., Parrotta, I., Burgio, F., Danesin, L., Busan, P., Zago, S., Mantini, D., D'Imperio, D., Zorzi, M., Filippini, N., Turolla, A., 10 2024. Modeling upper limb rehabilitation-induced recovery after stroke: The role of attention as a clinical confounder. *Physical Therapy* 105 (2), pzae148. URL: <https://doi.org/10.1093/ptj/pzae148> DOI: 10.1093/ptj/pzae148
- Sharir, O., Peleg, B., Shoham, Y., 2020. The cost of training nlp models: A concise overview. URL: <https://arxiv.org/abs/2004.08900>
- Turner, D. L., Ramos-Murguialday, A., Birbaumer, N., Hoffmann, U., Luft, A., 2013. Neurophysiology of robot-mediated training and therapy: a perspective for future use in clinical populations. *Frontiers in neurology* 4, 184.
- Wolf, S., Gerloff, C., Backhaus, W., 2021. Predictive value of upper extremity outcome measures after stroke—a systematic review and metaregression analysis. *Frontiers in Neurology* 12, 675255.
- Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., Wang, S., Yin, D., Du, M., 2024. Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology* 15 (2), 1–38.