



Guía de Asignatura

ASIGNATURA: *IA generativa: fundamentos y arquitecturas de modelos de lenguaje*

Título: *Máster de Formación Permanente en Inteligencia Artificial Aplicada a la Salud y Entornos Sanitarios*

Materia: *Fundamentos y tecnologías de IA en salud*

Créditos: 6 ECTS

Código: 05PASE

Índice

1.	Organización general	3
1.1.	Datos de la asignatura.....	3
1.2.	Introducción a la asignatura.....	3
1.3.	Competencias y resultados de aprendizaje	3
2.	Contenidos/temario	5
3.	Metodología	6
4.	Actividades formativas	7
5.	Evaluación.....	10
5.1.	Sistema de evaluación.....	10
5.2.	Directrices para el Uso de IA.....	10
5.3.	Sistema de calificación	11
6.	Bibliografía.....	12

1. Organización general

1.1. Datos de la asignatura

TITULACIÓN	<i>Máster de Formación Permanente en Inteligencia Artificial Aplicada a la Salud y Entornos Sanitarios</i>
ASIGNATURA	<i>IA generativa: fundamentos y arquitecturas de modelos de lenguaje grandes</i>
CÓDIGO - NOMBRE ASIGNATURA	<i>05PASE_IA generativa: fundamentos y arquitecturas de modelos de lenguaje grandes</i>
Carácter	Obligatorio
Cuatrimestre	Primero
Idioma en que se imparte	Castellano
Requisitos previos	No existen
Dedicación al estudio por ECTS	25 horas

1.2. Introducción a la asignatura

Esta asignatura proporciona la alfabetización necesaria para comprender los modelos de lenguaje grandes (LLMs) y su impacto estratégico en salud. Mediante el estudio de la arquitectura Transformer y el paradigma RAG, el estudiante adquiere las bases para liderar soluciones generativas con rigor técnico. El programa recorre desde el pre-entrenamiento hasta modelos especializados, permitiendo evaluar críticamente riesgos, alucinaciones y ética algorítmica. La materia capacita al egresado como interlocutor experto para traducir tecnología en valor clínico y seguridad del paciente. Estos fundamentos permiten liderar la transformación digital sanitaria, garantizando infraestructuras seguras y equidad en salud.

1.3. Competencias y resultados de aprendizaje

COMPETENCIAS GENERALES

CG.1.- Evaluar críticamente la aplicabilidad, limitaciones y riesgos de diferentes tecnologías de IA en contextos clínicos específicos, considerando tanto métricas técnicas como resultados clínicamente relevantes.

CG.2.- Diseñar e implementar estrategias de validación clínica de sistemas de IA, incluyendo la selección de diseños de estudio apropiados, definición de métricas de evaluación centradas en el paciente y protocolos de monitorización tras la implementación

COMPETENCIAS ESPECÍFICAS DE LA ASIGNATURA

C.E.1.- Analizar las arquitecturas y mecanismos de funcionamiento de los modelos de lenguaje grandes para evaluar su capacidad de razonamiento y generación en el dominio médico.

C.E.2.- Aplicar técnicas de recuperación aumentada de generación (RAG) y prompt engineering para optimizar la precisión, seguridad y trazabilidad de las respuestas de los modelos generativos en aplicaciones clínicas.

C.E.3.- Evaluar críticamente las limitaciones técnicas y los riesgos de alucinación o sesgo en el uso de IA generativa para garantizar la integridad, privacidad y equidad de la información en entornos sanitarios.

RESULTADOS DE APRENDIZAJE

Al finalizar esta asignatura se espera que el estudiante sea capaz de:

RA.1.- Categorizar las arquitecturas de los modelos de lenguaje (encoder-only, decoder-only e híbridos) y su adecuación técnica para diferentes tareas de procesamiento de lenguaje clínico y estructuración de datos.

RA.2.- Diferenciar el rendimiento de modelos de lenguaje generalistas frente a modelos especializados en el dominio biomédico mediante el análisis comparativo de benchmarks clínicos y casos de éxito validados.

RA.3.- Diseñar arquitecturas teóricas basadas en el paradigma RAG que permitan la integración de guías de práctica clínica y datos locales para mitigar la confabulación de datos en la toma de decisiones asistenciales.

2. Contenidos/temario

Tema 1. Evolución y paradigmas de la IA generativa

- ¿Qué es la IA generativa? Diferenciación entre IA discriminativa y generativa.
- Principios de pre-entrenamiento masivo (pre-training), ajuste fino (fine-tuning) y aprendizaje por refuerzo (RLHF).
- Escalado y ecosistema: leyes de escala (parámetros y datos) y análisis comparativo entre modelos open-source y propietarios.

Tema 2. La arquitectura Transformer y el mecanismo de atención

- Atención y lenguaje médico: impacto del paradigma self-attention para capturar dependencias semánticas en textos clínicos complejos.
- Componentes técnicos: gestión de la información mediante tokenización, embeddings, codificación posicional y capas de normalización.
- Variantes arquitectónicas: diferencias funcionales y casos de uso clínico de modelos encoder-only, decoder-only e híbridos.

Tema 3. Taxonomía de los Grandes Modelos de Lenguaje (LLMs)

- Familias GPT y BERT: evolución de arquitecturas autorregresivas y bidireccionales hacia la multimodalidad nativa.
- Modelos de vanguardia: capacidades de razonamiento clínico en familias LLaMA y Claude bajo criterios de seguridad constitucional.
- Eficiencia y multimodalidad: modelos de expertos e integración de texto, imagen y audio para medicina de precisión.

Tema 4. Modelos especializados en el dominio biomédico

- Adaptación biomédica: entrenamiento con corpus clínico y razonamiento médico avanzado en modelos de referencia como Med-PaLM 2.
- Especialización técnica: análisis de BioGPT, ClinicalBERT y modelos orientados a especialidades (radiología, patología y genómica).
- Evaluación y benchmarks: protocolos de validación mediante MedQA, USMLE y métricas de desempeño en el dominio médico.

Tema 5. Dinámicas cognitivas, limitaciones y arquitectura RAG

- Razonamiento y alucinaciones: análisis de procesos chain-of-thought y el origen técnico de la confabulación de datos en entornos sanitarios.
- Fundamentos de RAG: arquitectura de generación aumentada por recuperación para la trazabilidad de fuentes y mitigación de errores.
- Seguridad y ética: riesgos de memorización de datos, cumplimiento del RGPD y estrategias de mitigación de sesgos algorítmicos.

3. Metodología

La modalidad de enseñanza propuesta para el presente título guarda consonancia con la Metodología General de la Universidad Internacional de Valencia, aprobada por el Consejo de Gobierno Académico de la Universidad y de aplicación en todos sus títulos.

Este modelo, que vertebra el conjunto de procesos de enseñanza y aprendizaje de la institución, combina la naturaleza síncrona (mismo tiempo-diferente espacio) y asíncrona (diferente tiempo - diferente espacio) de los entornos virtuales de aprendizaje, siempre en el contexto de la modalidad virtual.

El elemento síncrono se materializa en sesiones de diferente tipo (clases expositivas y prácticas, tutorías, seminarios y actividades de diferente índole durante las clases online) donde el profesor y el estudiante comparten un espacio virtual y un tiempo determinado que el estudiante conoce con antelación.

Las actividades síncronas forman parte de las actividades formativas necesarias para el desarrollo de la asignatura y, además, quedan grabadas y alojadas para su posterior visualización.

Por otro lado, estas sesiones no solamente proporcionan espacios de encuentro entre estudiante y profesor, sino que permiten fomentar el aprendizaje colaborativo, al generarse grupos de trabajo entre los estudiantes en las propias sesiones.

Los elementos asíncronos del modelo se desarrollan a través del Campus Virtual, que contiene las aulas virtuales de cada asignatura, donde se encuentran los recursos y contenidos necesarios para el desarrollo de actividades asíncronas, así como para la interacción y comunicación con los profesores y con el resto de departamentos de la Universidad.

4. Actividades formativas

La metodología VIU, basada en la modalidad virtual, se concreta en una serie de actividades formativas y metodologías docentes que articulan el trabajo del estudiante y la docencia impartida por los profesores.

Durante el desarrollo de cada una de las asignaturas, se programan una serie de actividades de aprendizaje que ayudan a los estudiantes a consolidar los conocimientos trabajados en cada una de las asignaturas. A continuación, listamos las actividades genéricas que pueden formar parte de cada asignatura, dependiendo de las competencias a desarrollar en los estudiantes en cada asignatura.

1. Clases presenciales

2. Clases virtuales síncronas

Constituyen el conjunto de acciones formativas que ponen en contacto al estudiante con el profesor, con otros expertos y con compañeros de la misma asignatura en el mismo momento temporal a través de herramientas virtuales. Las actividades recurrentes (por ejemplo, las clases) se programan en el calendario académico y las que son ocasionales (por ejemplo, sesiones con expertos externos) se avisan mediante el tablón de anuncios del campus. Estas actividades se desglosan en las siguientes categorías:

a. Clases expositivas: El profesor expone a los estudiantes los fundamentos teóricos de la asignatura.

b. Clases prácticas: El profesor desarrolla junto con los estudiantes actividades prácticas que se basan en los fundamentos vistos en las clases expositivas. En términos generales, su desarrollo consta de las siguientes fases, pudiéndose adaptar en función de las necesidades docentes:

I. La primera fase se desarrolla en la sala principal de la videoconferencia, donde el profesor plantea la actividad.

II. A continuación, divide a los estudiantes en grupos de trabajo a través de las salas colaborativas y se comienza con la actividad. En esta fase el profesor va entrando en cada sala colaborativa rotando los grupos para resolver dudas, dirigir el trabajo o dar el feedback oportuno. Los estudiantes también tienen posibilidad de consultar al profesor en el momento que consideren necesario.

III. La tercera fase también se desarrolla en la sala principal y tiene como objetivo mostrar el ejercicio o explicar con ejemplos los resultados obtenidos. Por último, se ponen en común las conclusiones de la actividad realizada.

No obstante, el profesor puede utilizar otras metodologías activas y/o herramientas de trabajo colaborativo en estas clases.

c. Seminarios: En estas sesiones un experto externo a la Universidad acude a presentar algún contenido teórico-práctico directamente vinculado con el temario de la asignatura. Estas sesiones permiten acercar al estudiante a la realidad de la disciplina en términos no sólo profesionales, sino también académicos. Todas estas sesiones están vinculadas a contenidos de las asignaturas y del programa educativo.

3. Actividades asíncronas supervisadas

Se trata de un conjunto de actividades supervisadas por el profesor de la asignatura vinculadas con la adquisición por parte de los estudiantes de los resultados de aprendizaje y el desarrollo de sus competencias. Estas actividades, diseñadas con visión de conjunto, están relacionadas entre sí para ofrecer al estudiante una formación completa e integral. Esta categoría se desglosa en el siguiente conjunto de actividades:

a. Actividades y trabajos prácticos: se trata de un conjunto de actividades prácticas realizadas por el estudiante por indicación del profesor que permiten al estudiante adquirir las competencias del título, especialmente aquellas de carácter práctico. Estas actividades, entre otras, pueden ser de la siguiente naturaleza: actividades vinculadas a las clases prácticas (resúmenes, mapas conceptuales, one minute paper, resolución de problemas, análisis reflexivos, generación de contenido multimedia, exposiciones de trabajos, test de autoevaluación, participación en foros, entre otros). Estas actividades serán seleccionadas por el profesor en función de las necesidades docentes. Posteriormente, estas actividades son revisadas por el profesor, que traslada un feedback al estudiante sobre las mismas, pudiendo formar parte de la evaluación continua de la asignatura.

b. Actividades guiadas con recursos didácticos audiovisuales e interactivos: se trata de un conjunto de actividades en las que el estudiante revisa o emplea recursos didácticos (bibliografía, videos, recursos interactivos) bajo las indicaciones realizadas previamente por el profesor; con el objetivo de profundizar en los contenidos abordados en las sesiones teóricas y prácticas. Estas sesiones permiten la reflexión o práctica por parte del estudiante, y pueden complementarse a través de la puesta en común en clases síncronas o con la realización de actividades y trabajos prácticos. Posteriormente, estas actividades son revisadas por el profesor, que traslada un feedback al estudiante sobre las mismas, pudiendo formar parte de la evaluación continua de la asignatura.

4. Tutorías

En esta actividad se engloban las sesiones virtuales de carácter síncrono y las comunicaciones por correo electrónico o campus virtual destinadas a la tutorización de los estudiantes. En ellas, el profesor comparte información sobre el progreso del trabajo del estudiante a partir de las evidencias recogidas, se resuelven dudas y se dan orientaciones específicas ante dificultades concretas en el desarrollo de la asignatura. Pueden ser individuales o colectivas, según las necesidades de los estudiantes y el carácter de las dudas y orientaciones planteadas. Tal y como se ha indicado, se realizan a través de videoconferencia y e-mail.

Se computan una serie de horas estimadas, pues, aunque existen sesiones comunes para todos los estudiantes, éstos posteriormente pueden solicitar al docente tantas tutorías como estimen necesarias.

Dado el carácter mixto de esta actividad formativa, se computa un porcentaje de sincronía estimado del 30%.

5. Estudio autónomo

En esta actividad el estudiante consulta, analiza y estudia los manuales, bibliografía y recursos propios de la asignatura de forma autónoma a fin de lograr un aprendizaje significativo y superar la evaluación de la asignatura de la asignatura. Esta actividad es indispensable para adquirir las competencias del título, apoyándose en el aprendizaje autónomo como complemento a las clases y actividades supervisadas.

6. Examen final

Como parte de la evaluación de cada una de las asignaturas (a excepción de las prácticas y el Trabajo fin de título), se realiza una prueba o examen final. Esta prueba se realiza en tiempo real (con los medios de control antifraude especificados) y tiene como objetivo evidenciar el nivel de adquisición de conocimientos y desarrollo de competencias por parte de los estudiantes. Los exámenes o pruebas de evaluación final se realizan en las fechas y horas programadas con antelación y con los sistemas de vigilancia online (proctoring) de la universidad.

5. Evaluación

5.1. Sistema de evaluación

El Modelo de Evaluación de estudiantes en la Universidad se sustenta en los principios del Espacio Europeo de Educación Superior (EEES), y está adaptado a la estructura de formación virtual propia de esta Universidad. De este modo, se dirige a la evaluación de competencias.

Sistema de Evaluación	Ponderación
Portafolio*	60 %
Sistema de Evaluación	Ponderación
Prueba final*	40 %

***Es requisito indispensable para superar la asignatura aprobar cada apartado (portafolio y prueba final)** con un mínimo de 5 para ponderar las calificaciones.

Los enunciados y especificaciones propias de las distintas actividades serán aportados por el docente, a través del Campus Virtual, a lo largo de la impartición de la asignatura.

Atendiendo a la Normativa de Evaluación de la Universidad, se tendrá en cuenta que la utilización de **contenido de autoría ajena** al propio estudiante debe ser citada adecuadamente en los trabajos entregados. Los casos de plagio serán sancionados con suspenso (0) de la actividad en la que se detecte. Asimismo, el uso de **medios fraudulentos durante las pruebas de evaluación** implicará un suspenso (0) y podrá implicar la apertura de un expediente disciplinario.

5.2. Directrices para el Uso de IA

1. **Declaración obligatoria.** Si utiliza una herramienta de IA para la elaboración de cualquier trabajo evaluable, debe incluir una declaración al final del mismo. Esta declaración debe ser clara y detallada, especificando:
 - **Herramienta utilizada.** Nombre completo y versión de la herramienta (ej., ChatGPT 4.0, Midjourney 6.0).
 - **Propósito.** La razón específica por la que utilizó la IA (ej., "para generar un borrador", "para buscar bibliografía", "para corregir la sintaxis").
 - **Impacto en el resultado.** Cómo influyó en el trabajo final (ej., "el borrador generado sirvió como base para la introducción, pero fue reescrito y ampliado por completo").
 - **Productos integrados.** Los fragmentos o elementos generados por la IA que se incorporaron directamente en la entrega (ej., "la tabla 2 fue generada por IA y validada con datos de la OMS").
2. **Responsabilidad total del estudiante.** La validez, corrección, originalidad y ética de todo el trabajo recaen exclusivamente en el/la estudiante. No basta con generar contenido con IA; es su responsabilidad revisarlo, validarlo y adaptarlo para que refleje tu esfuerzo y

conocimiento. Esto incluye la verificación de datos, la corrección de errores y la reelaboración del texto para que tenga su propia voz y estilo.

3. **Límites y plagio.** Se considerará **plagio** y una falta grave de integridad académica el presentar como propio contenido generado por IA sin la debida reelaboración y validación. Esto también aplica si el uso de la IA tiene como objetivo eludir el aprendizaje real.

Consecuencias del incumplimiento

Cualquier caso de plagio o fraude académico detectado resultará en un **suspenso en la asignatura**. El objetivo de esta medida es salvaguardar la calidad del proceso de enseñanza-aprendizaje y garantizar que la evaluación refleje el conocimiento y las habilidades que has adquirido.

5.3. Sistema de calificación

La calificación de la asignatura se establecerá en los siguientes cálculos y términos:

Nivel de aprendizaje	Calificación numérica	Calificación cualitativa
Muy competente	9,0 - 10	Sobresaliente
Competente	7,0 - 8,9	Notable
Aceptable	5,0 - 6,9	Aprobado
Aún no competente	0,0 - 4,9	Suspenso

Sin detrimento de lo anterior, el estudiante dispondrá de una **rúbrica simplificada** en el aula que mostrará los aspectos que valorará el docente, como así también los **niveles de desempeño que tendrá en cuenta para calificar las actividades vinculadas a cada resultado de aprendizaje**.

La mención de «**Matrícula de Honor**» podrá ser otorgada a estudiantes que hayan obtenido una calificación igual o superior a 9. Su número no podrá exceder del cinco por ciento de los estudiantes matriculados en una materia en el correspondiente curso académico, salvo que el número de estudiantes matriculados sea inferior a 20, en cuyo caso se podrá conceder una sola «Matrícula de Honor».

6. Bibliografía

- Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, 183(6), 589. <https://doi.org/10.1001/jamainternmed.2023.1838>
- Bhuyan, S. S., Sateesh, V., Mukul, N., Galvankar, A., Mahmood, A., Nauman, M., Rai, A., Bordoloi, K., Basu, U., & Samuel, J. (2025). Generative artificial intelligence use in healthcare: Opportunities for clinical excellence and administrative efficiency. *Journal of Medical Systems*, 49(1), 10. <https://doi.org/10.1007/s10916-024-02136-1>
- Brodsky, V., Ullah, E., Bychkov, A., Song, A. H., Walk, E. E., Louis, P., Rasool, G., Singh, R. S., Mahmood, F., Bui, M. M., & Parwani, A. V. (2025). Generative artificial intelligence in anatomic pathology. *Archives of Pathology & Laboratory Medicine*, 149(4), 298-318. <https://doi.org/10.5858/arpa.2024-0215-RA>
- Chen, D., Alnassar, S. A., Avison, K. E., Huang, R. S., & Raman, S. (2025). Large language model applications for health information extraction in oncology: Scoping review. *JMIR Cancer*, 11, e65984. <https://doi.org/10.2196/65984>
- Chen, D., Parsa, R., Hope, A., Hannon, B., Mak, E., Eng, L., Liu, F.-F., Fallah-Rad, N., Heesters, A. M., & Raman, S. (2024). Physician and artificial intelligence chatbot responses to cancer questions from social media. *JAMA Oncology*. <https://doi.org/10.1001/jamaoncol.2024.0836>
- Cuadra, A., Wang, M., Stein, L. A., Jung, M. F., Dell, N., Estrin, D., & Landay, J. A. (2024). The illusion of empathy? Notes on displays of emotion in human-computer interaction. *Proceedings of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3613904.3642336>
- Fleurence, R. L., Bian, J., Wang, X., Xu, H., Dawoud, D., Higashi, M., & Chhatwal, J. (2024). Generative artificial intelligence for health technology assessment: Opportunities, challenges, and policy considerations: An ISPOR working group report. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 28(2), 175-183. <https://doi.org/10.1016/j.ival.2024.10.3846>
- Habicht, J., Viswanathan, S., Carrington, B., Hauser, T. U., Harper, R., & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature Medicine*, 30(2), 595–602. <https://doi.org/10.1038/s41591-023-02766-x>
- Huo, B., Boyle, A., Marfo, N., Tangamornsuksan, W., Steen, J. P., McKechnie, T., Lee, Y., Mayol, J., Antoniou, S. A., Thirunavukarasu, A. J., Sanger, S., Ramji, K., & Guyatt, G. (2025). Large language models for chatbot health advice studies: A systematic review. *JAMA Network Open*, 8(2), e2457879. <https://doi.org/10.1001/jamanetworkopen.2024.57879>
- Iqbal, U., Tanweer, A., Rahmanti, A. R., Greenfield, D., Lee, L. T., & Li, Y. J. (2025). Impact of large language model (ChatGPT) in healthcare: An umbrella review and evidence synthesis. *Journal of Biomedical Science*, 32(1), 45. <https://doi.org/10.1186/s12929-025-01131-z>

- Kim, K., Cho, K., Jang, R., Kyung, S., Lee, S., Ham, S., Choi, E., Hong, G. S., & Kim, N. (2024). Updated primer on generative artificial intelligence and large language models in medical imaging for medical professionals. *Korean Journal of Radiology*, 25(3), 224-242. <https://doi.org/10.3348/kjr.2023.0818>
- Kolding, S., Lundin, R. M., Hansen, L., & Østergaard, S. D. (2024). Use of generative artificial intelligence (AI) in psychiatry and mental health care: A systematic review. *Acta Neuropsychiatrica*, 37, e37. <https://doi.org/10.1017/neu.2024.50>
- Li, C., Zhang, Y., Weng, Y., Wang, B., & Li, Z. (2023). Natural language processing applications for computer-aided diagnosis in oncology. *Diagnostics (Basel, Switzerland)*, 13(2). <https://doi.org/10.3390/diagnostics13020286>
- Lu, M. Y., Chen, B., Williamson, D. F. K., Chen, R. J., Zhao, M., Chow, A. K., Ikemura, K., Kim, A., Pouli, D., Patel, A., Soliman, A., Chen, C., Ding, T., Wang, J. J., Gerber, G., Liang, I., Le, L. P., Parwani, A. V., Weishaupt, L. L., & Mahmood, F. (2024). A multimodal generative AI copilot for human pathology. *Nature*, 634(8033), 466-473. <https://doi.org/10.1038/s41586-024-07618-3>
- Montemayor, C., Halpern, J., & Fairweather, A. (2022). In principle obstacles for empathic AI: Why we can't replace human empathy in healthcare. *AI & Society*, 37(4), 1353–1359. <https://doi.org/10.1007/s00146-021-01230-z>
- Saab, K., Tu, T., Weng, W.-H., Tanno, R., Stutz, D., Wulczyn, E., Zhang, F., Strother, T., Park, C., Vedadi, E., Chaves, J. Z., Hu, S.-Y., Schaeckermann, M., Kamath, A., Cheng, Y., Barrett, D. G. T., Cheung, C., Mustafa, B., Palepu, A., ... Natarajan, V. (2024). Capabilities of Gemini models in medicine. *arXiv*. <https://doi.org/10.48550/arXiv.2404.18416>
- Shanmugam, D., Agrawal, M., Movva, R., Chen, I. Y., Ghassemi, M., Jacobs, M., & Pierson, E. (2025). Generative artificial intelligence in medicine. *Annual Review of Biomedical Data Science*, 8(1), 199-226. <https://doi.org/10.1146/annurev-biodatasci-103123-095332>
- Shool, S., Adimi, S., Saboori Amleshi, R., Bitaraf, E., Golpira, R., & Tara, M. (2025). A systematic review of large language model (LLM) evaluations in clinical medicine. *BMC Medical Informatics and Decision Making*, 25(1), 117. <https://doi.org/10.1186/s12911-025-02954-4>
- Thangaraj, P. M., Benson, S. H., Oikonomou, E. K., Asselbergs, F. W., & Khera, R. (2024). Cardiovascular care with digital twin technology in the era of generative artificial intelligence. *European Heart Journal*, 45(45), 4808-4821. <https://doi.org/10.1093/eurheartj/ehae619>
- Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine*, 29(8), 1930-1940. <https://doi.org/10.1038/s41591-023-02448-8>
- Van Veen, D., Van Uden, C., Blankemeier, L., Delbrouck, J. B., Aali, A., Bluethgen, C., Pareek, A., Polacin, M., Reis, E. P., Seehofnerová, A., Rohatgi, N., Hosamani, P., Collins, W., Ahuja, N., Langlotz, C. P., Hom, J., Gatidis, S., Pauly, J., & Chaudhari, A. S. (2024). Adapted large language models can outperform medical experts in clinical text summarization. *Nature Medicine*, 30(4), 1134-1142. <https://doi.org/10.1038/s41591-024-02855-5>

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf