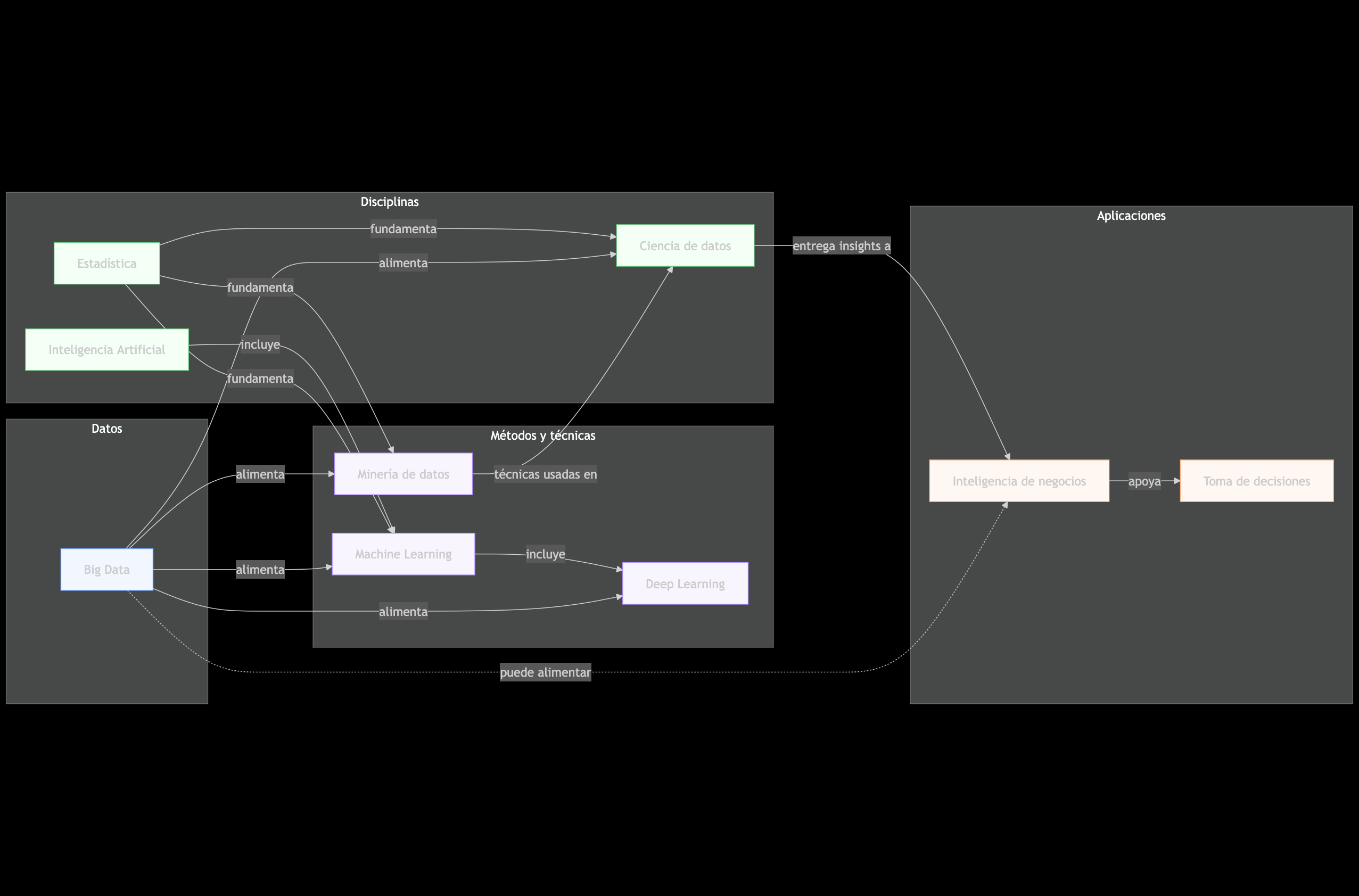
Nombre: Jhon Henry Quintero

Taller 1 - Introducción a Sistemas Inteligentes

1. Utilice un modelo grande de lenguaje (LLM) —tal como ChatGPT, Bard o Bing— para obtener una descripción breve de los siguientes términos:
   1. **Big Data:** Conjunto de datos tan grandes, variados y veloces que no pueden ser procesados eficazmente con métodos tradicionales. Se caracteriza por las “3V”: volumen, variedad y velocidad.
   2. **Machine Learning (Aprendizaje automático):** Subcampo de la inteligencia artificial que permite a las computadoras aprender de los datos y mejorar su desempeño en tareas específicas sin ser programadas explícitamente.
   3. **Inteligencia Artificial (IA):** Área de la informática que busca desarrollar sistemas capaces de realizar tareas que normalmente requieren inteligencia humana, como reconocer imágenes, comprender lenguaje o tomar decisiones.
   4. **Ciencia de datos (Data Science):** Disciplina que combina estadística, matemáticas, programación y conocimiento del negocio para extraer información útil y generar valor a partir de los datos.
   5. **Deep Learning (Aprendizaje profundo):** Rama del machine learning que utiliza redes neuronales artificiales con muchas capas para modelar patrones complejos, especialmente en imágenes, voz y texto.
   6. **Minería de datos (Data Mining):** Proceso de descubrir patrones, relaciones y tendencias en grandes volúmenes de datos mediante técnicas estadísticas, de aprendizaje automático y de bases de datos.
   7. **Inteligencia de negocios (Business Intelligence, BI):** Conjunto de estrategias, tecnologías y herramientas que transforman los datos de una empresa en información procesable para apoyar la toma de decisiones.
   8. **Estadística:** Ciencia que recopila, organiza, analiza e interpreta datos para describir fenómenos y tomar decisiones basadas en evidencia.
2. Utilice un LLM para crear un diagrama que muestre las relaciones entre los anteriores conceptos. Pídale que genere el diagrama usando la sintaxis de mermaid.js y visualícelo usando https://mermaid.live/. Opcionalmente puede usar un plugin. Discuta el resultado, ¿tiene sentido? Pruebe dándole instrucciones al LLM para que lo mejore.  
     
   
3. Busque un video reciente sobre alguna aplicación o técnica de inteligencia artificial que usted considere muy interesante. Véalo y haga una síntesis. Escoja un segmento de unos 30 segundos para mostrar al resto de la clase.
   1. <https://youtu.be/zXiUXIskduM> (IA para detección de cáncer de seno)
4. Use un LLM para determinar qué son los transformers. Busque información al respecto sobre los siguientes aspectos:

**a) Ideas principales de los transformers**

* Usan **autoatención** (self-attention) para ponderar qué tokens del contexto importan.
* **Multi-head attention**, **residuals** y **positional encodings** estabilizan y capturan orden.
* Estructuras típicas: **encoder–decoder** (traducción) y **decoder-only** (generación).
* Se entrenan en paralelo (sin recurrencias), lo que los hace **rápidos de entrenar**.

**b) Aplicaciones**

* **NLP**: traducción, resumen, chat, búsqueda, QA.
* **Visión**: clasificación/detección (Vision Transformers).
* **Multimodal**: texto-imagen (p. ej., captioning, CLIP), audio/voz.
* **Código**: asistencia de programación y generación de funciones.

**c) ¿Qué es un LLM y su relación con transformers?**

* Un **LLM** es (casi siempre) un **transformer decoder-only** grande.
* Objetivo base: **predecir el siguiente token** sobre enormes corpus.
* Por eso, los LLMs son **una aplicación a gran escala** de la arquitectura transformer.

**d) ¿Cómo se entrena un LLM conversacional?**

1. **Pre-entrenamiento** autoregresivo (next-token) en texto masivo.
2. **SFT**: afinado supervisado con pares instrucción→respuesta.
3. **Alineación por preferencias** (p. ej., **RLHF** o **DPO**) para respuestas útiles y seguras.
4. **Guardas y evaluación**: filtros, pruebas de seguridad y mejoras continuas.
5. Utilice un LLM para resolver algunos de los problemas de programación en el curso:

<https://uecp.edunext.io/courses/course-v1:uecp+mlds0a+001/about>

1. Busque tareas de cursos que haya tomado recientemente que se puedan resolver con un LLM, pruebe el LLM y discuta los resultados.
   1. Para el curso de procesos modernos de desarrollo de sw, desarrollamos un wiki enteramente usando IA, <https://github.com/jvillada1/Procesos-Modernos-de-DLLO>, para ello usamos Groq y los agentes disponibles para generar resultados a diferentes preguntas relacionadas con minecraft
2. Use el LLM para encontrar argumentos a favor y en contra de la posibilidad de una inteligencia artificial general.
   * **A favor**
     1. **No hay impedimentos físicos conocidos**: si la mente humana es un proceso computable, en principio puede ser emulada por máquinas.
     2. **Precedente biológico**: la inteligencia general existe (humanos); por tanto, es alcanzable en algún sustrato.
     3. **Escalado funciona**: al aumentar datos, parámetros y cómputo surgen capacidades nuevas (razonamiento básico, código, multitarea).
     4. **Sistemas compuestos**: LLM + herramientas externas (búsqueda, memoria, planificación, actuadores) amplían la “generalidad”.
     5. **Mejora recursiva asistida**: la propia IA ya acelera I+D (programación, experimentación), lo que podría impulsar avances.
     6. **Eficiencia algorítmica e ingeniería**: mejores arquitecturas, entrenamiento y hardware reducen requisitos para nuevas capacidades.
   * **En contra**
     1. **Definición resbaladiza**: “AGI” es ambiguo; mover la meta hace imposible “alcanzarla” de forma consensuada.
     2. **Generalización y causalidad**: los modelos actuales fallan fuera de distribución, en razonamiento causal y en planes de largo horizonte.
     3. **Símbolos sin anclaje**: sin experiencia sensorimotora, el “significado” puede quedar débil (problema de *grounding*).
     4. **Datos y sostenibilidad**: el texto público es finito; entrenar al nivel actual requiere datos/cómputo/energía enormes.
     5. **Fiabilidad**: alucinaciones, falta de calibración y robustez minan la “generalidad” práctica.
     6. **Límites computacionales**: muchas tareas generales (planificación óptima, búsqueda exhaustiva) son intratables en el peor caso.
     7. **Riesgos y fricción externa**: seguridad, regulación y costes podrían frenar o reconducir el desarrollo antes de llegar a AGI.

8. Use el LLM para indagar sobre los riesgos e implicaciones éticas de la inteligencia artificial

**Riesgos principales**

* **Sesgos y discriminación:** modelos entrenados con datos sesgados replican o amplifican injusticias.
* **Privacidad y vigilancia:** recolección masiva de datos, reidentificación y uso sin consentimiento.
* **Desinformación y deepfakes:** amplificación de noticias falsas y manipulación de la opinión pública.
* **Opacidad y falta de rendición de cuentas:** decisiones difíciles de explicar o auditar.
* **Seguridad y abuso:** uso malicioso (phishing, automatización de ciberataques), fallos no intencionales.
* **Impacto laboral y desigualdad:** automatización que desplaza tareas y concentra poder económico.
* **Sostenibilidad ambiental:** altos costos energéticos y de cómputo.
* **Militarización y control social:** uso en armas autónomas o para reprimir libertades.

**Implicaciones éticas**

* **Justicia y equidad**, **respeto a la autonomía y la dignidad**, **protección de derechos**, **distribución justa de beneficios y riesgos**, y **gobernanza responsable**.

Link repo: <https://github.com/mr-sudaca/especializacion-sw-dev/tree/main/semestre1/ml/taller1>