Dyania Health Synapsis AI vs ChatGPT

Large Language Model (LLM) applications in Clinical Research

OpenAI's natural language processing (NLP) tool Chat GPT has garnered increasing public attention as one of the most well-known **Large Language Models (LLMs)**. The health sciences industry remains one of the last industries to seek adoption of this new technological advancement, and it is viewed with both skepticism and excitement amongst medical professionals. Education is needed to debunk media misconceptions about ChatGPT, clarify the capabilities and limitations of LLMs, and demonstrate how Synapsis AI makes LLMs useful for medical sciences.

For an industry-wide adoption of this new technology to occur, medical professionals and NLP applied scientists must **bridge the knowledge gap between Medicine and A.I.** through mutual education, understanding, and awareness. The most important facet of this goal is to establish an **in-depth understanding of what LLMs can do and what LLMs cannot do, specifically as it relates to medicine**. Mutual education fosters realistic expectations and satisfaction with due to fit-for-purpose applications.

LLMs can predict the insights and data to be drawn from free text but alone do not suffice for clinical research applications. They cannot infer or reason to solve a problem or question asked of the data. For completeness in human-like critical thinking, the questions asked of the data must be algorithmically developed with a logic based on medical standard-of-care protocols as well as current research by physicians and medical professionals.

Synapsis AI Bridges Medicine & Mathematics

>30 Billion parameter LLM for drawing appropriate conclusions from full electronic medical records at scale with a temporal awareness for order and duration of clinical events

Clinical logic designed by board certified MD's and PharmD's for algorithmic inference and matching clinical conclusions and insights with clinical research protocol criteria

DYANIA HEALTH

What is an LLM?

LLMs are statistical models that predict a missing word from a sentence after being trained on (learning from) huge volumes of text. LLMs do not ”think” like humans do but instead predict the likely next word, words, or phrases. For example, when posing the question “What is the capital of the United Kingdom?”, the LLM predicts the most likely next word which is “London”. While LLMs can often get the answers so incredibly correct that they appear intelligent, they are only designed and trained to predict the missing words, not think critically, infer, or draw conclusions.

The text data commonly used to train LLMs is from a data-set called “The Pile”. The Pile includes every available text on the internet ranging from news articles to Jane Austen novels. A notable point is that The Pile does not include electronic medical records or medical school textbooks, but it does include PubMed data.

Large language models are ”large” because they are comprised of billions of nodes that are connected by parameters. The models are designed and inspired by the way synapses exist between neurons in the human brain, sending signals from one to another. These webs of nodes take inputs, send signals from one to another, and spit out an output. If the goal was to determine whether a paragraph is describing a patient who took a certain therapy an LLM ”reads” the text and predicts the output that answers the question or completes the missing words.

Key Limitations of ChatGPT

- **Error-Prone:** illogical output or “hallucinations”
- **Biased:** skewed towards common data point probabilities (or only data available on the internet)
- **Limited Text Size Capacity:** cannot process large amounts of text or combine information from various text sources
- **Inconsistent:** Results are not reproducible
- **HIPAA Non-Compliant:** lacking patient privacy safeguards
- **Outdated:** most recent data cutoff was in 2021
- **Lacking Logic and Inference:** cannot do reasoning or even simple arithmetic
- **Lack of Interpretability:** impossible to track source of information or justifications

See Table 1 for more details

DYANIA HEALTH

How LLMs are trained and Why it Matters in Clinical Research

1 Pre-Training

Pre-training a model with billions of words in text is focused on **generally training** the model to predict words in a sentence. One way to do this is by pre-training a Masked Language Model where bi-directional models (like BERT by Google). In this case, a percent of the words in the training data are masked and the model's task is to predict the missing words. The model can see words before and after the missing words, in other words in both directions, making it a bi-directional model. Another kind of model pre-training is auto-regressive (like GPT by OpenAI) which are uni-directional and trained to predict the next word without seeing succeeding words. This results in these types of models being better performing for language generation. Overall, an LLM is only pre-trained once and this is currently an expensive and time-consuming process.

2 Fine-Tuning

Fine-tuning adds **a task-specific layer** to the LLM where the model parameters are updated through gradient descent on a hyper-specific dataset. We can also freeze certain layers and fine-tune the rest to get better performance.

3 Instruction-Training

Instruction training is performed to train the models to follow **specific types of instructions** with human feedback. This can be done through in-context learning through prompts. In this case instead of billions of texts like those used in the pre-training stage, the model takes a few task-specific examples as input and then understands the context of the task before making a prediction using this information.

Overall, for LLMs to be effective at drawing conclusions from electronic medical records and clinical research, the fine-tuning and instruction training needs to be specific to the medical applications. In other words, LLM's can only be as effective in medicine as the humans who fine-tune and instruction-train them. **At Dyania Health, we have a team of physicians and pharmacists completing fine-tuning and instruction training to our Synapsis AI models.**

Unlike ChatGPT, Dyania Health pre-adapts, instruction-tunes and fine-tunes our Synapsis AI models on a huge quantity of extremely high-quality and clean medical, pharmaceutical, and biological knowledge, including several proprietary datasets compiled by our own clinicians, including >250,000 physician notes. These models are then used to attain an uncommonly deep understanding of patient EMRs and draw conclusions about them, particularly in connection with eligibility criteria for clinical trials and observational retrospective clinical studies.

Natural Language Processing Comparison – Large Language Models	
ChatGPT	**Synapsis AI**
Error-Prone: ChatGPT within healthcare has been known to fail at basic math, make up research sources, and "hallucinate" - providing realistic yet factually incorrect responses (Forbes: Apr 2023) known as "hallucinations".	**Medically Sound Logic:** Synapsis AI contains pre-optimized clinical rules to algorithmically match conclusions or interpretations from electronic medical records with clinical research criteria. Criteria are structured as machine-friendly and objective queries against information found in free text and structured data form (labs, codes, etc.).
Biased: ChatGPT responses are based on probabilities of availability of data, meaning, clinical indications that are not common will likely be factually incorrect. This is particularly problematic for studying rare diseases and medically underserved populations.	**Unbiased:** Because Synapsis AI is focused on the medical domain, we can afford to take more proactive and effective steps to de-bias our models and counteract existing systemic discrimination tendencies. Synapsis AI identifies all eligible candidates for a given clinical trial and devotes particular effort to analyzing and flagging the eligibility of patients from traditionally disadvantaged backgrounds.
Inconsistent: ChatGPT does not rely on custom ontologies or logical structures to interpret text, many responses are generated at random and, therefore, outputs are difficult to reproduce.	**Reliable:** The application of clinical rules enables Synapsis AI to synthesize accurate and logical clinical conclusions ready for physician review.
HIPAA Non-Compliant: In its current state, there is no way to use ChatGPT with PHI that would be considered HIPAA compliant due to their terms of use which clearly state the utilization of log data, device information, and usage data including IP addresses, content engaged with, actions taken, dates & times used, etc.	**Patient Data Security:** Synapsis AI is deployed behind the healthcare systems' firewall, meaning no data leaves the healthcare systems' IT computing network. All data processing is conducted on a sandbox server that is given read-only access to clinically relevant data necessary for the specific targeted studies.
Outdated: While ChatGPT's model has recently been updated, the data pile is restricted to 2021 - requiring systematic and frequent retraining to keep pace with new publications, biomarkers, drug indications, etc.	**Bleeding edge:** Trained on latest EMR data, the model has been trained on hundreds of thousands of medical records, both common and outlier indications, inclusive of outcomes data with all evidence being traceable back to the source of truth.

Dyania Health is a venture-backed company founded in 2020 which has developed Synapsis AI, a fit-for-purpose natural language processing tool for EMR-based clinical research. Dyania was founded by a seasoned entrepreneurial team, co-founded by a senior medical advisory board of P.I.'s at academic centers across the US. The engineering team is led by senior ex-Amazon Alexa NLP scientists who work alongside an in-house physician team developing and ever-improving Synapsis AI's clinical decision engine.