

Key Takeaways:

1. **Limitations of Large Context Windows:** Large context windows in LLMs, despite their potential, do not necessarily facilitate deep synthesis of long documents. Decomposing documents into atomic reasoning units remains crucial for extracting meaningful insights.
2. **Systems of Systems Approach:** Brightwave leverages multiple subsystems to address specific tasks within the larger model. This includes using rag and agent-based reasoning to draw conclusions from disparate data sources, illustrating the importance of integrating various machine learning subsystems for holistic analysis.
3. **Domain-Specific Chunking Strategies:** The effectiveness of chunking strategies, particularly within financial documents, is highlighted. Recognizing and modeling the intent of specific sections of documents (e.g., S1 filings) helps in accurate information retrieval and minimizes overlooking critical information.
4. **Factuality Assurance Techniques:** Brightwave employs several innovative methods for ensuring factual accuracy, including multiple passes over material, generating consensus from various model iterations, and training dedicated models to verify that generated insights are supported by primary sources.
5. **Integrating Domain Expertise:** The necessity of integrating domain expertise into model development and deployment is emphasized. Brightwave's success hinges on understanding the specifics of financial documents and strategically guiding the AI's focus during retrieval and inference processes.

Follow-up Ideas:

1. **Experiment with Modular Systems:** Implement a modular approach for a financial analysis LLM by developing separate models for chunking, information retrieval, and inference. Integrate these modules to assess how effectively they can work together to provide comprehensive analysis.
2. **Develop and Validate Chunking Algorithms:** Design and test chunking algorithms tailored to financial documents, such as S1 filings or earnings reports. Conduct experiments to evaluate how different chunking strategies impact the retrieval and synthesis of relevant financial information.