

Present and Future: Current Trends in LLMs



Multimodal

Text, image, audio, video integration



Mixture of Experts

Sparse activation for efficiency



Long Context

100K+ tokens processing



Better Reasoning

Math, logic, commonsense



Retrieval Augmentation

Combining LMs with knowledge bases for accurate information



Smaller Efficient Models

On-device and edge deployment capabilities



Enhanced Capabilities

Improvements in mathematical reasoning, logical thinking, and commonsense understanding



Context Processing

Processing 100K+ tokens enables handling entire books and documents



Specialized Domain Models

Medical

Legal

Scientific

Finance

Engineering



AGI Considerations

Focus on **capabilities**, **safety**, and **alignment** as we approach AGI