

ประโยชน์ของ RLHF



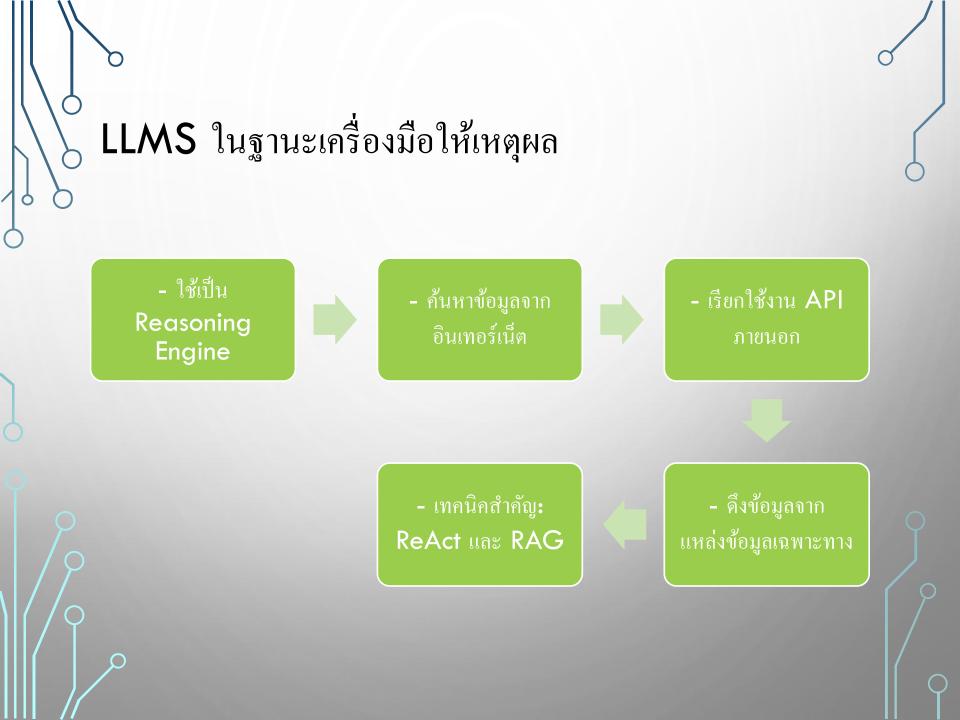
- ลดการสร้างเนื้อหาที่ไม่พึงประสงค์



- ทำให้โมเคลมีความซื่อสัตย์, มีประโยชน์, และปลอดภัย



- นักวิจัยยังคงพัฒนาเทคนิคนี้เพื่อให้ Al มีความรับผิดชอบมากขึ้น



สรุปสั้น ๆ

- RLHF ช่วย
ปรับแต่งโมเดลให้
มีความปลอดภัย
และเป็นประโยชน์

- Responsible
Al เป็นหัวข้อ
สำคัญที่นักวิจัย
กำลังให้ความสนใจ

- LLMs สามารถ เรียก API และ ฐานข้อมูลได้

- เทคนิคสำกัญ: ReAct และ RAG

LLM-POWERED APPLICATIONS



1. DISTILLATION (การกลั่นความรู้)



2. QUANTIZATION (การลดความละเอียดของ ตัวเลข)



3. PRUNING (การตัดแต่งโมเคล)



- ลบพารามิเตอร์ที่ไม่มีผลต่อผลลัพธ์มาก



- Structured Pruning: ลบโหนด/เลเยอร์ที่ไม่สำคัญ



- Unstructured Pruning: ลบน้ำหนักที่มีค่าต่ำ



- ลดขนาดของโมเคล ทำให้ inference เร็วขึ้น