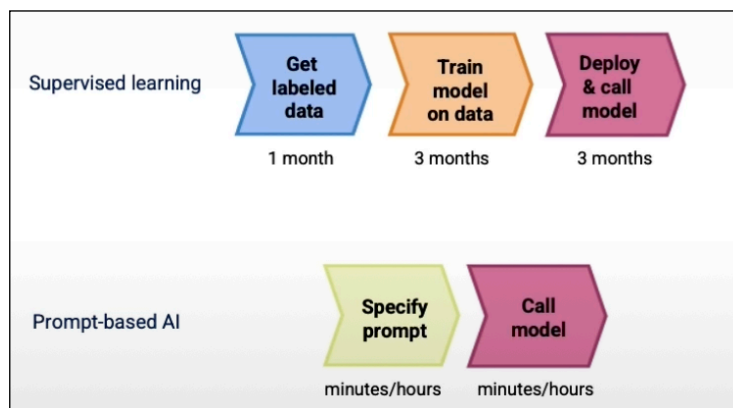


## How Prompting will Revolutionize Automation Workflows in Finance



Andrew Ng

### **Objective:**

1. Processing tasks for chain-of-thought & multi step reasoning with financial information
2. Showcase the best practices when building systems using a LLM

### **Initial Intention**

Replicate the prompting-based method described in the paper linked with the dataset. Additionally, I aim to replicate the error scenarios outlined in the paper.

### **Tools & Packages Used:**

Python 3.11.5, Jupyter Notebook, OpenAI, Langchain

### **Challenges & Limitations:**

1. An application often needs multiple internal steps that are invisible to the end user. How to clearly understand those internal steps?
2. How to define a threshold for number of input tokens to the model?

### **Future Work:**

1. Building a complex multi step bot application for numerical reasoning on financial reports
2. Look into the ways of improving the system over time

### **Reference:**

1. Paper **“CONVFINQA: Exploring the Chain of Numerical Reasoning in Conversational Finance Question Answering”**

The study introduces the CONVFINQA dataset, aiming to simulate human reasoning in real-world finance contexts. Experimental findings reveal that neural symbolic models, despite using tailored architectures and large scale data, fall short of human expert performances. Prompting-based few-shot learning, while leveraging memory recall with prompts, struggles with new task paradigms and complex calculations, indicating challenges in adapting to diverse domains. The study suggests encoding various task paradigms into large language models for clearer reasoning, but acknowledges the need for domain-specific models for effective performance. Moreover, due to cost constraints, experiments are limited to the GPT-3 model, leaving room for further investigation into advanced prompt engineering and larger pre-trained language models. The findings highlight the potential of prompting-based methods while underscoring the necessity for future research to address the challenges of complex reasoning tasks in specific domains.