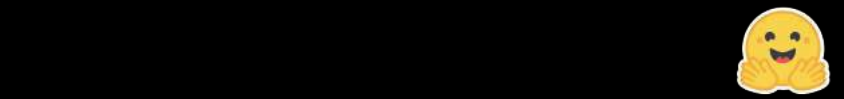
Enterprise & LLMs

March 2023

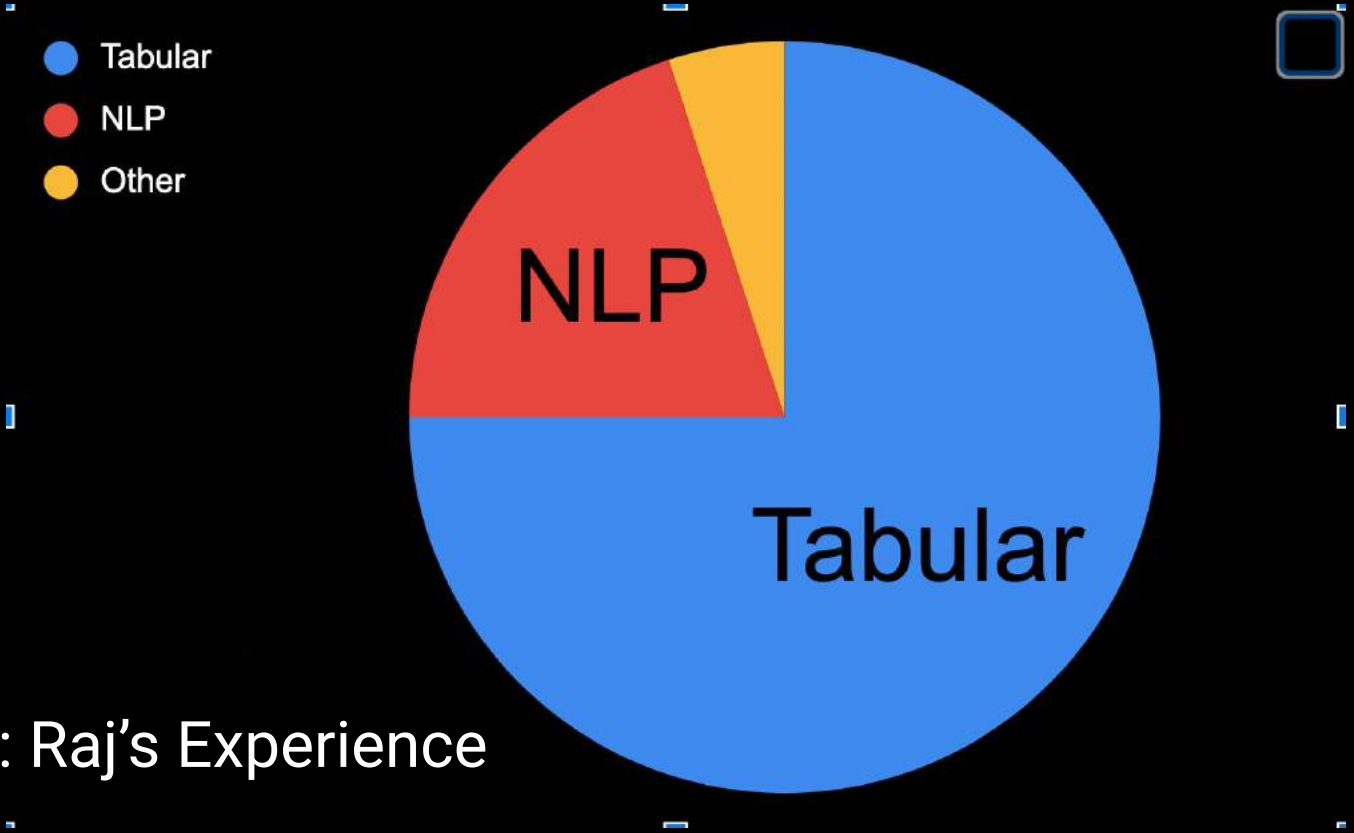
Rajiv Shah | @rajistics
raj@hf.co

Overview

- NLP in the Enterprise
- Effect of LLMs
  - Existing NLP use case
  - New opportunities
- What you should do today



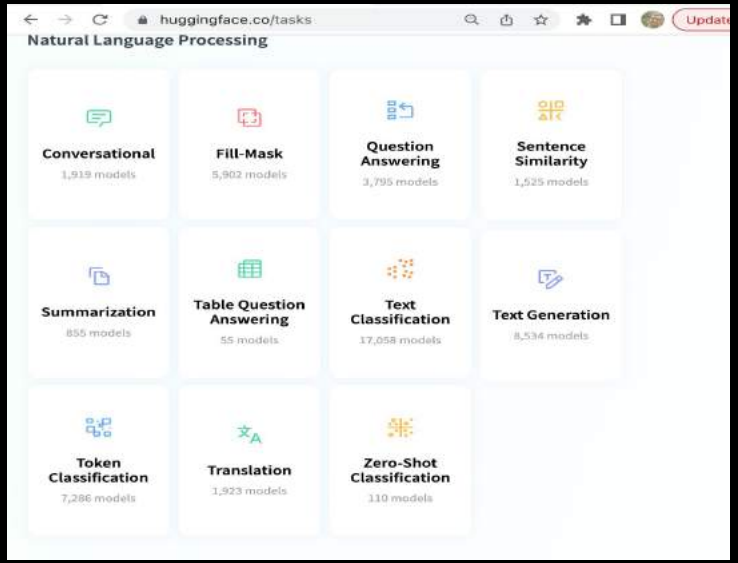
Machine Learning Use Cases in the Enterprise



Source: Raj’s Experience

Will LLMs Displace Traditional NLP Models?

Ranking Interest in NLP Use Cases



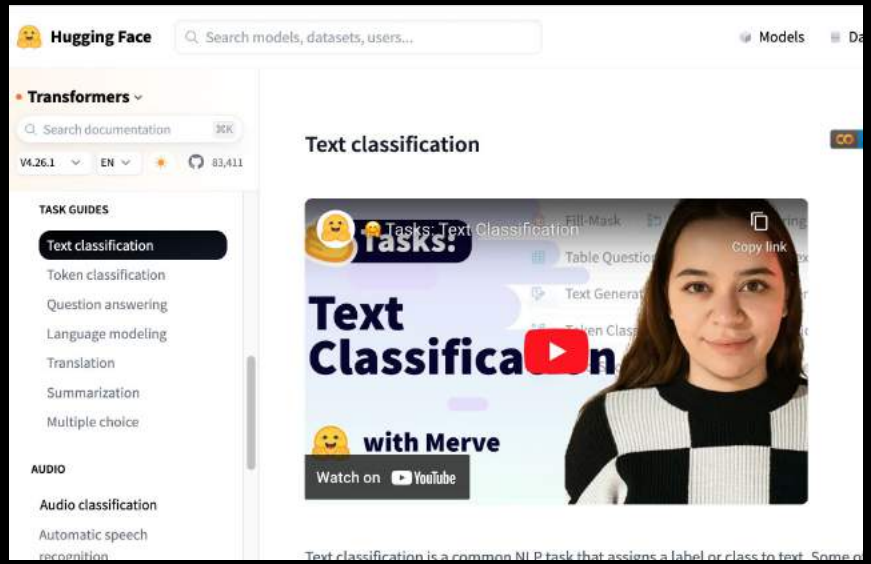
Tasks

1. Text Generation
2. Question Answering
3. Text Classification
4. Summarization
5. Sentence Similarity

Analytics

Ranking Interest in NLP Use Cases

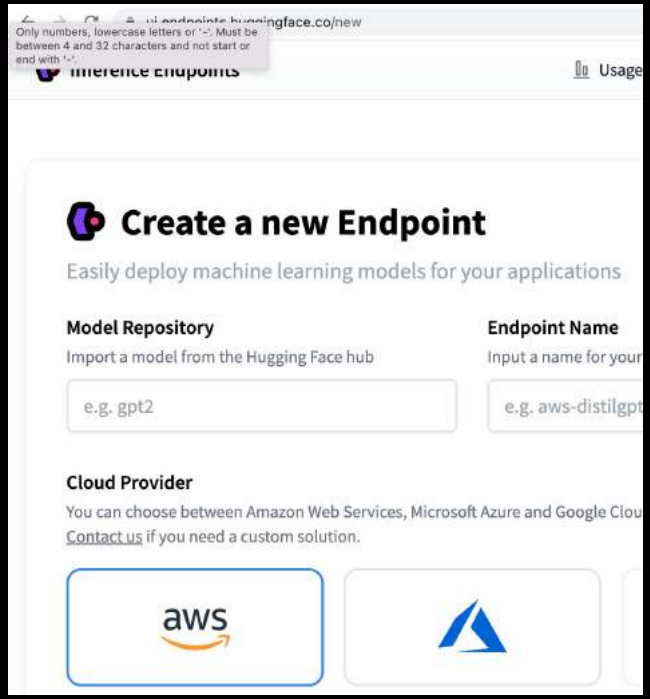
Docs



1. Text Classification
2. Summarization
3. Question Answering
4. Token Classification

Analytics

Ranking NLP deployments



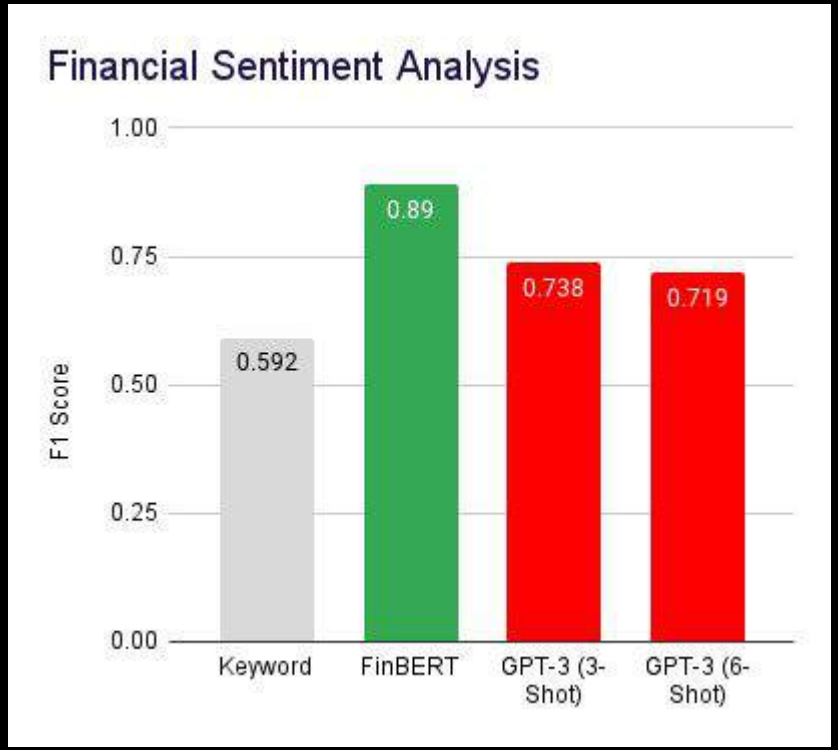
1. **Text Classification**
2. Sentence Embeddings
3. Text Generation
4. Token Classification
5. Question Answering
6. Summarization

Source: Hugging Face Deployment Stats

Let’s Compare LLMs to Traditional NLP Models

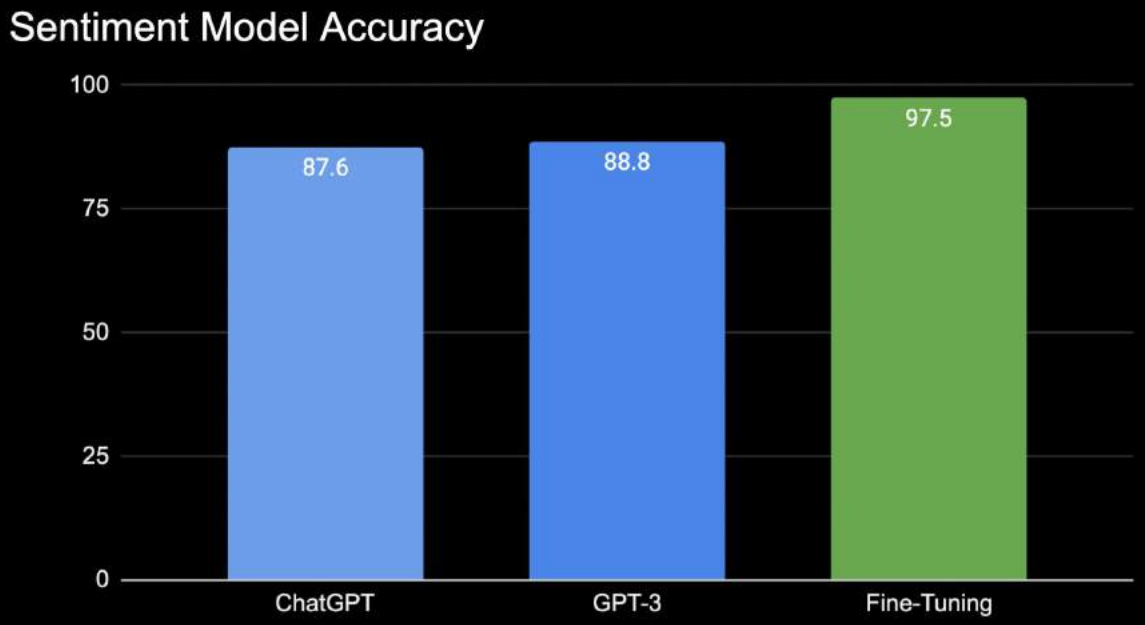
LLMs for Text Classification

- Financial domain
- Over a 15% Drop in Accuracy



https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4337182

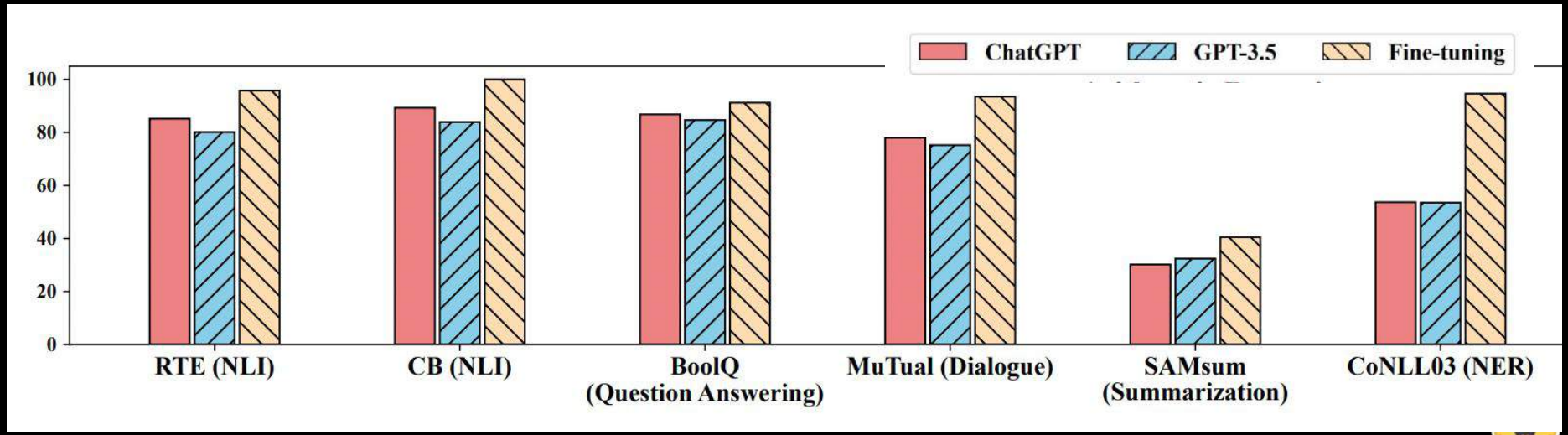
LLMs for Text Classification



https://arxiv.org/pdf/2302.06476.pdf

LLMs across common NLP tasks

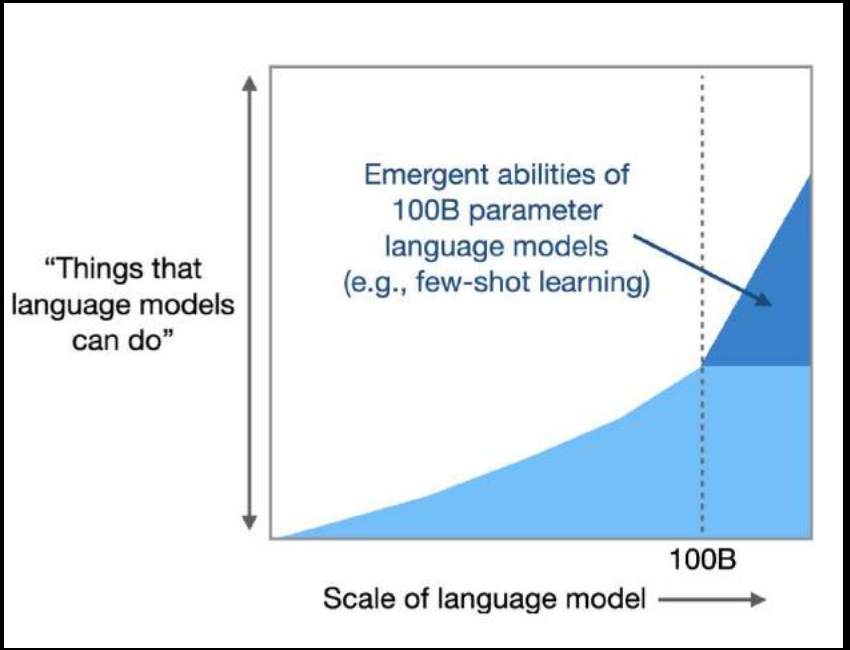
- Zero Shot versus Fine-Tuned Models



https://arxiv.org/pdf/2302.06476.pdf

Few Shot Learning

LLMs are intriguing because of the emergence of In-Context Learning



Jason Wei

Few Shot Learning

- Give the model a few examples
- Then use it to predict in the same way

Given a news article, classify its topic.
Possible labels: 1. World 2. Sports 3. Business 4. Sci/Tech

Article: A nearby star thought to harbor comets and asteroids now appears to be home to planets, too.
Label: Sci/Tech

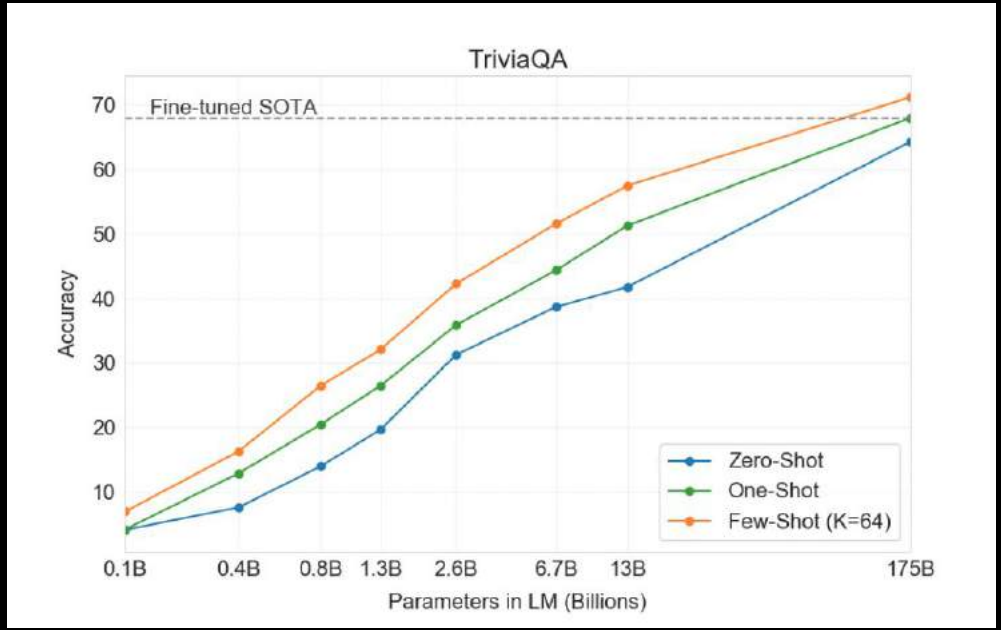
Article: Soaring crude prices plus worries about the economy and the outlook for earnings are expected to hang over the stock market next week during the depth of the summer doldrums.
Label: Business

Article: Murtagh a stickler for success Northeastern field hockey coach Cheryl Murtagh doesn't want the glare of the spotlight that shines on her to detract from a team that has been the America East champion for the past three years and has been to the NCAA tournament 13 times.
Label:

Few-Shot Learning - GPT-3

- GPT-3:

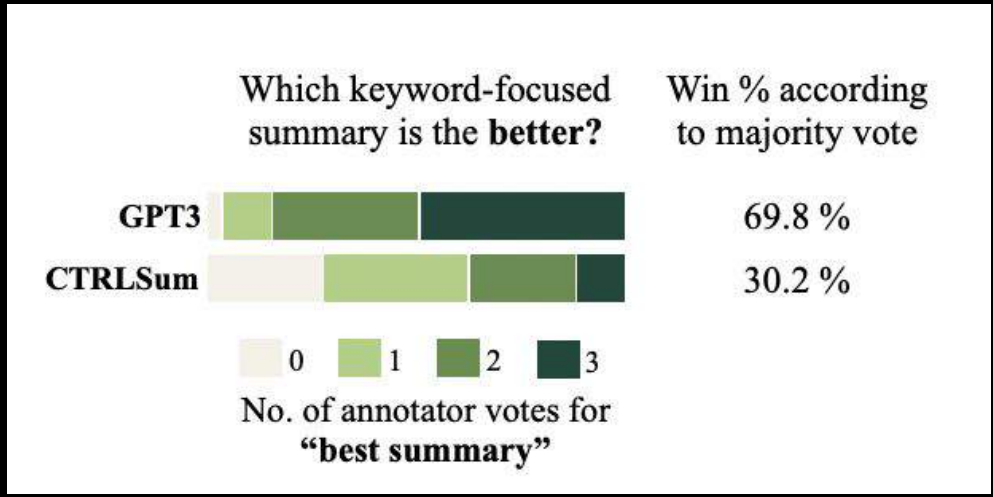
  nearly matching
  the performance
  of state-of-the-art
  fine-tuned systems



https://arxiv.org/pdf/2005.14165.pdf

Few-Shot Learning - News

News Summarization: GPT-3 is better than fine-tuning & close to human level

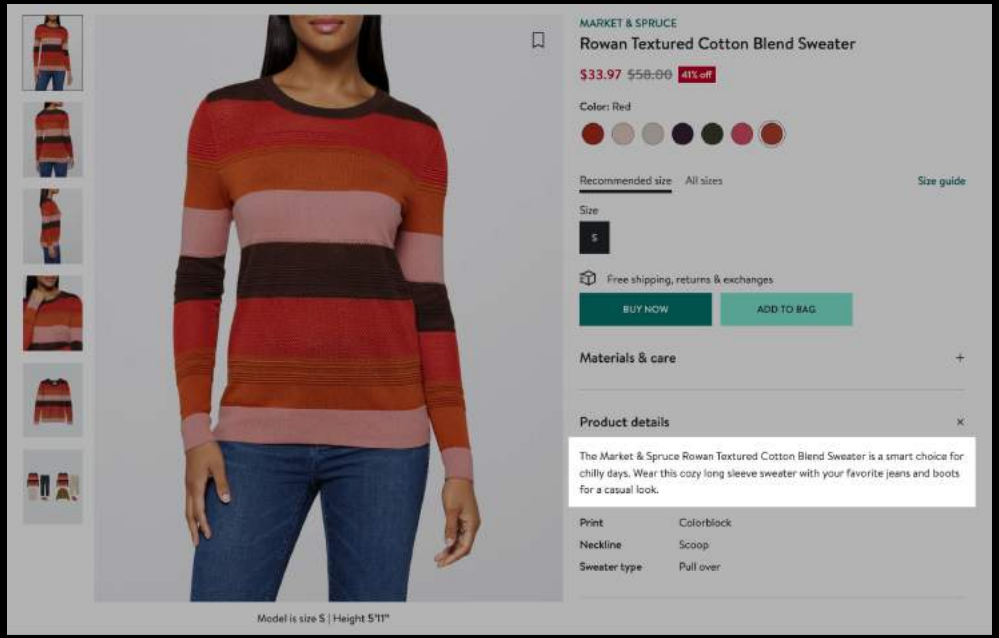


https://arxiv.org/pdf/2209.12356.pdf
https://arxiv.org/pdf/2301.13848.pdf

Fine Tuned LLMs

Product Descriptions:

AI Generated descriptions from several hundred fine tuned examples were rated higher than human written examples



https://multithreaded.stitchfix.com/blog/2023/03/06/expert-in-the-loop-generative-ai-at-stitch-fix/

How do Large Language Models Do?

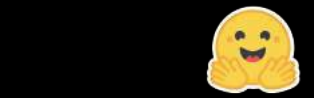
When to use LLMs?

For domain specific tasks -> Fine Tuning wins

For high scale/sensitive applications -> Dedicated open source model

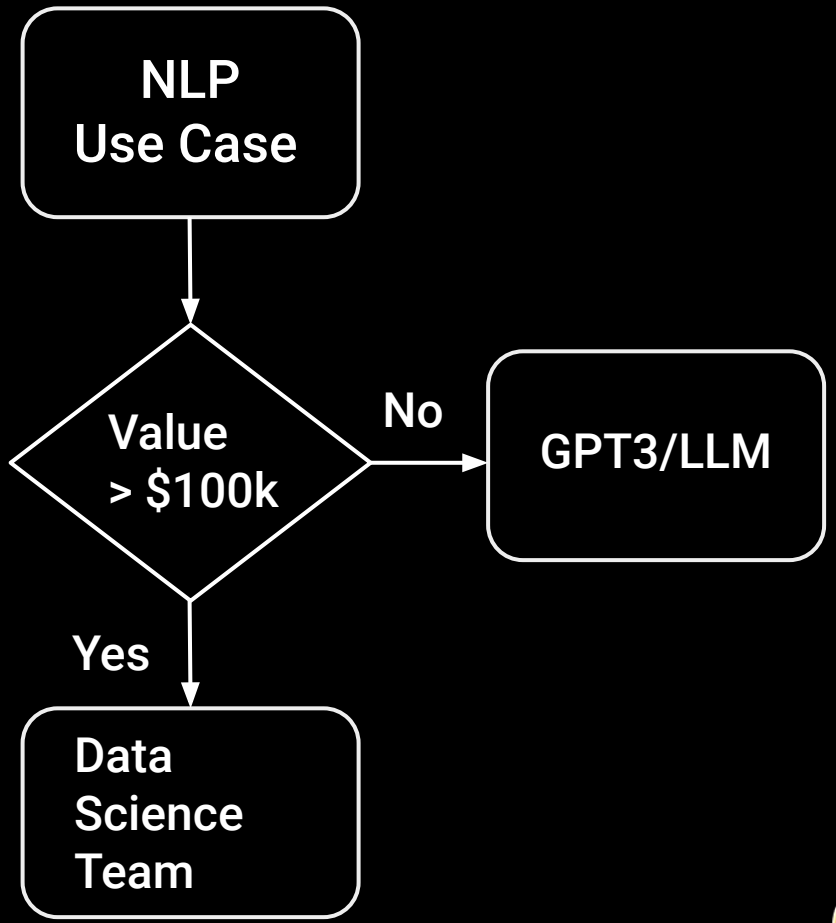
For many general tasks:

- Few shot performance is close
- Zero shot performance is still good
- This gap is likely to close quickly



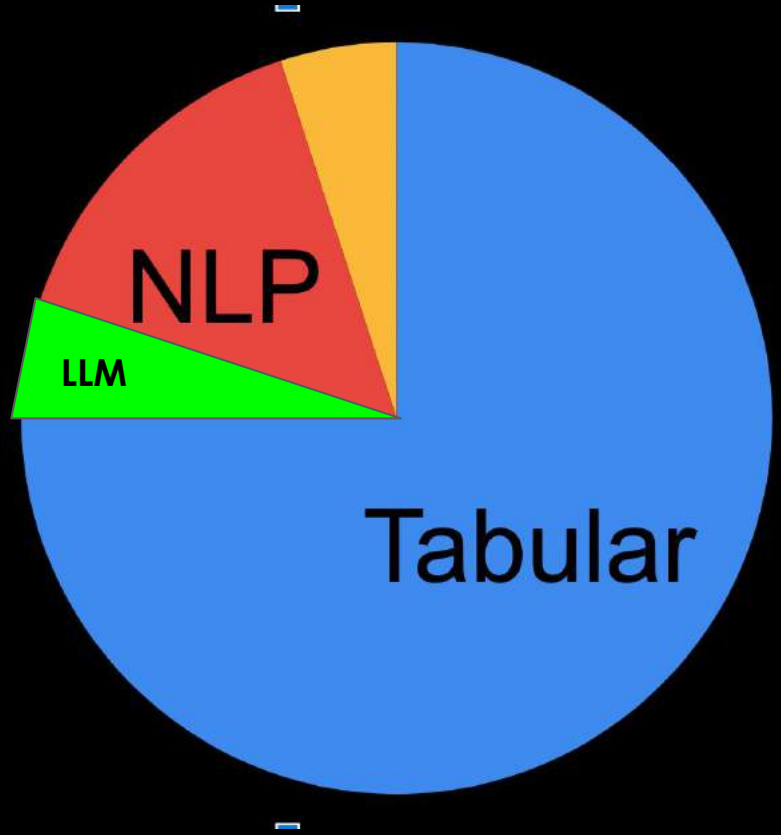
Impact of LLMs on Existing NLP Use Cases

Encourage using LLMs by *educated* users for the low value use cases.



Impact of LLMs on Existing NLP Use Cases

LLMs will likely have low impact on the current enterprise analytics roadmap




User Perspective

ChatGPT Sprints to One Million Users

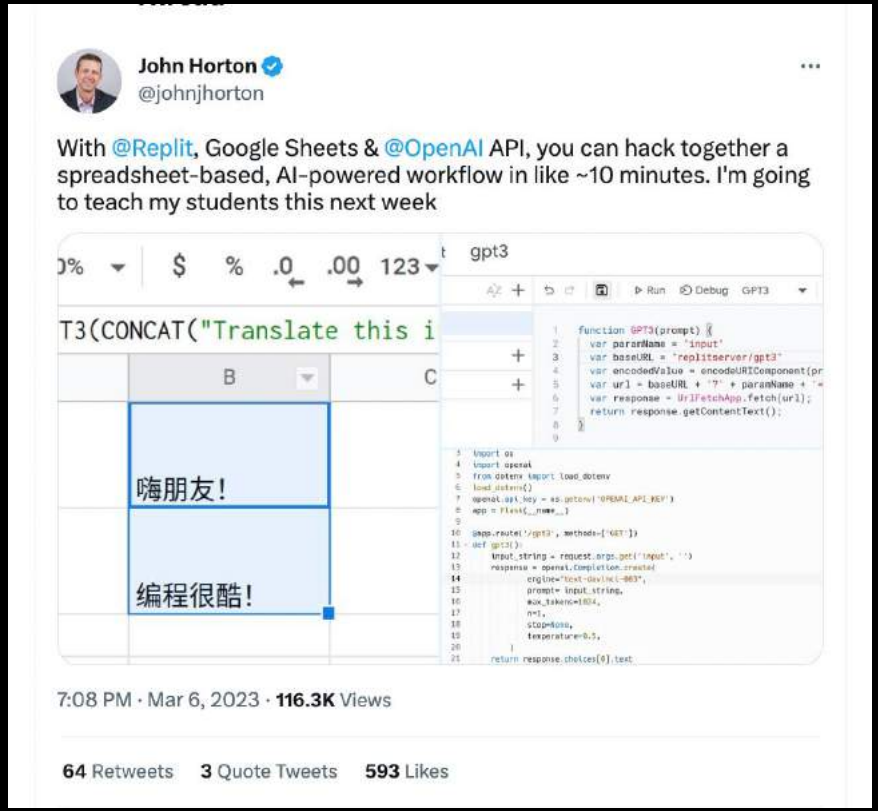
Time it took for selected online services to reach one million users

* one million backers ** one million nights booked *** one million downloads

Source: Company announcements via Business Insider/Linkedin

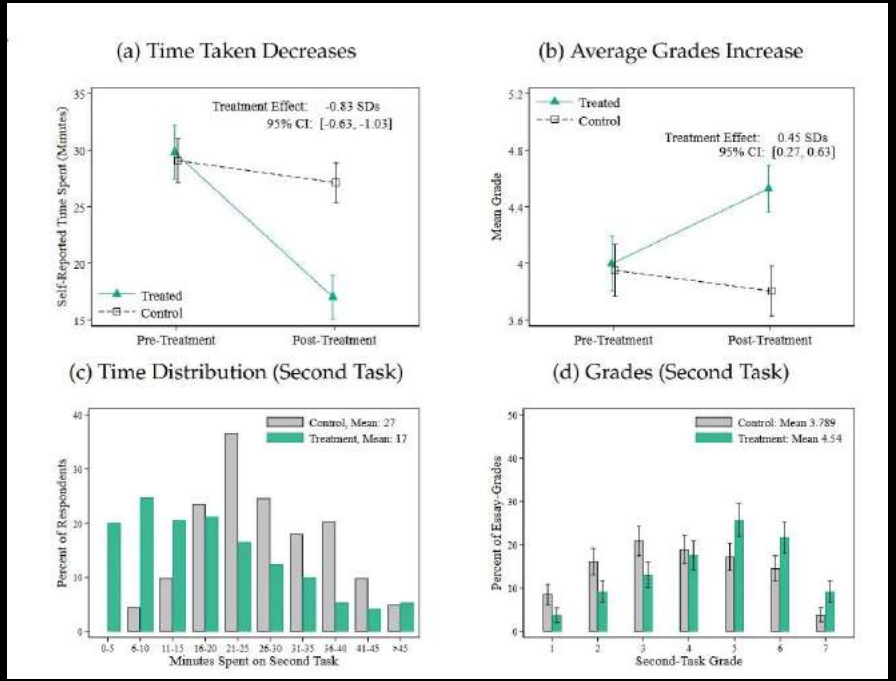
statista

Lot of Enthusiasm for using LLMs



Using LLMs

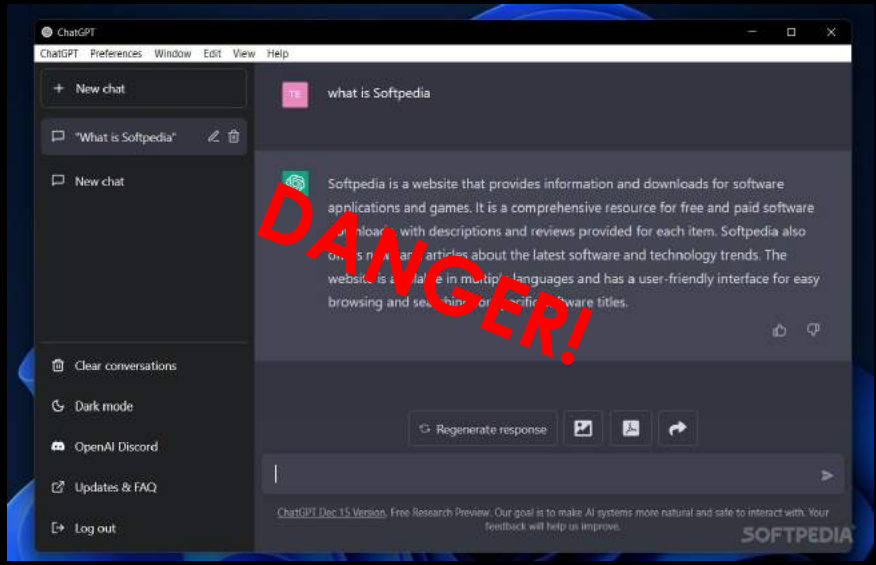
Raises productivity for mid-level professional writing tasks



https://economics.mit.edu/sites/default/files/inline-files/Noy_Zhang_1.pdf

People like to use LLMs for answers

- People use it like a search engine - they like the quick answers
- **Reinvigorating Search**



Proper Way to use LLMs for Search

- Information Retrieval + LLM
- 🗂️ LlamaIndex 🦙 (GPT Index)
  - Central interface to connect your LLM's with external data.

https://github.com/jerryjliu/gpt_index

LLMs as Decision Makers

- LLMs can make intermediate decisions (Agent)
  - Intent classification
- LLMs can generate text to connect to APIs and services

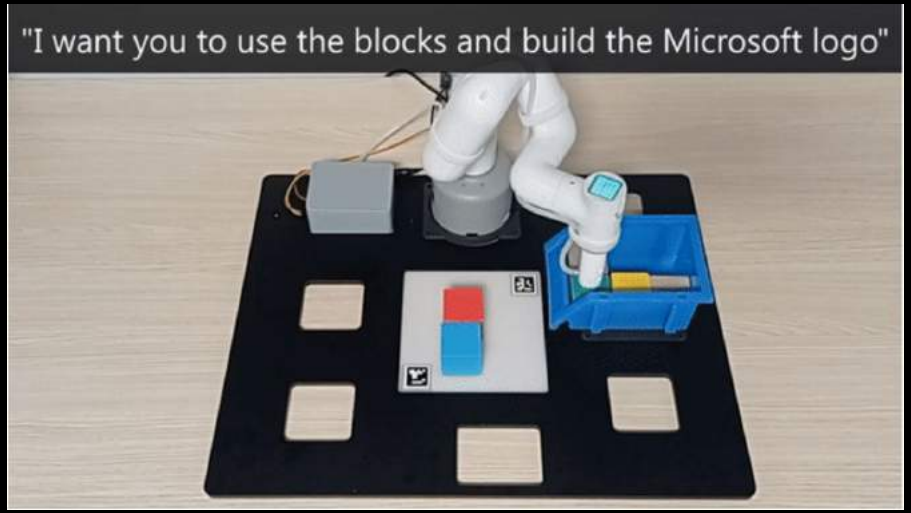
```
> Entering new AgentExecutor chain...
 I need to find out who Leo DiCaprio's girlfriend
is and then calculate her age raised to the 0.43
power.
Action: Search
Action Input: "Who is Leo DiCaprio's girlfriend?"
Observation: Camila Morrone
Thought: I need to find out Camila Morrone's age
Action: Search
Action Input: "How old is Camila Morrone?"
Observation: 25 years
Thought: I need to calculate 25 raised to the
0.43 power
Action: Calculator
Action Input: 25^0.43

> Entering new LLMMathChain chain...
25^0.43
```python
import math
print(math.pow(25, 0.43))
```
```

https://langchain.readthedocs.io/en/latest/modules/agents/implementations/mrkl.html

LLMs Can Work with Many Diverse Services

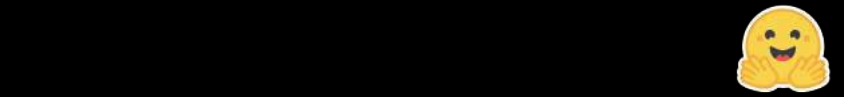
- Extending ChatGPT to control robots



https://www.microsoft.com/en-us/research/group/autonomous-systems-group-robotics/articles/chatgpt-for-robotics/

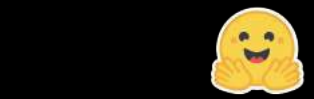
Recommendations for Enterprise Analytics

- **Start building knowledge of using LLMs**
  - Building prompting skills / LangChain
  - Start educating the rest of the organization on how to use LLMs properly
    - Identify the crucial use cases
    - Start collecting data on the interactions your users have
  - Prepare for multiple LLM landscape (OpenAI, Cohere, Hugging Face, . . .)



Recommendations for Enterprise Analytics

- **Use this excitement to prioritize search related projects.**
  - Start building Information Retrieval systems
  - Have that prompt engineer integrating multiple systems
- **Start looking at other services/API that can interconnect with LLMs**



Enterprise & LLMs

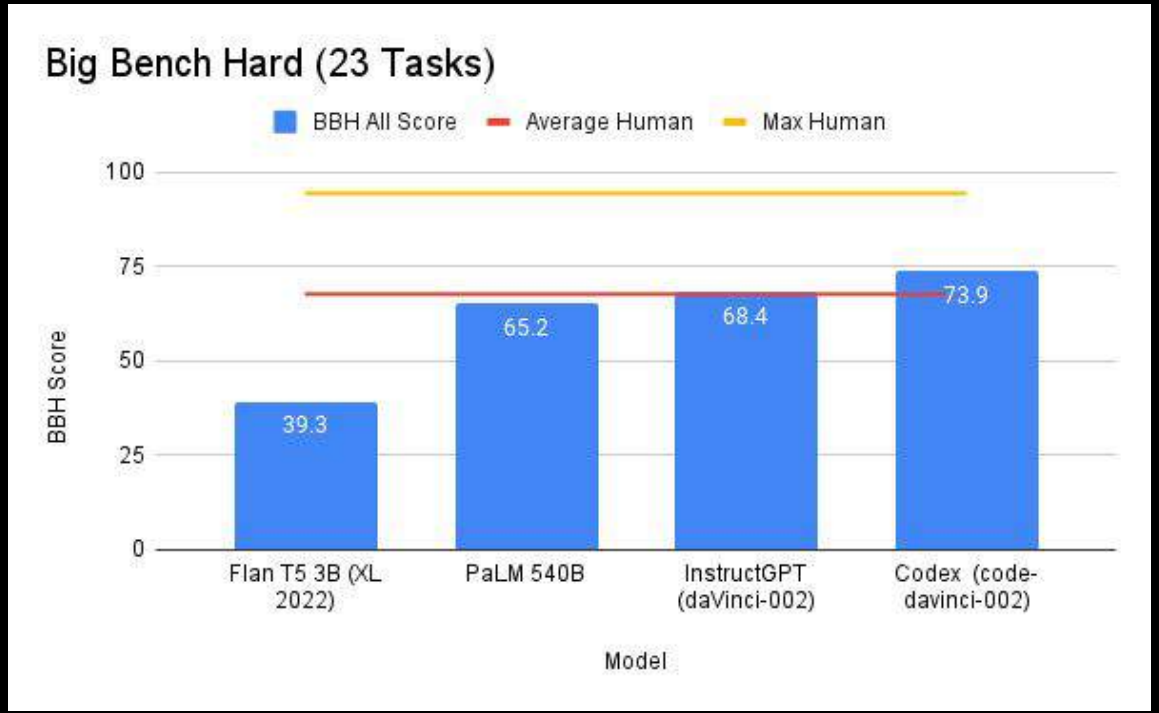
March 2023

Rajiv Shah | @rajistics
raj@hf.co

Ability for LLMs to Perform Reasoning Tasks

Very good at solving reasoning tasks:

- Algorithmic/Arithmetic
- Natural Language Understanding
- Use of World Knowledge



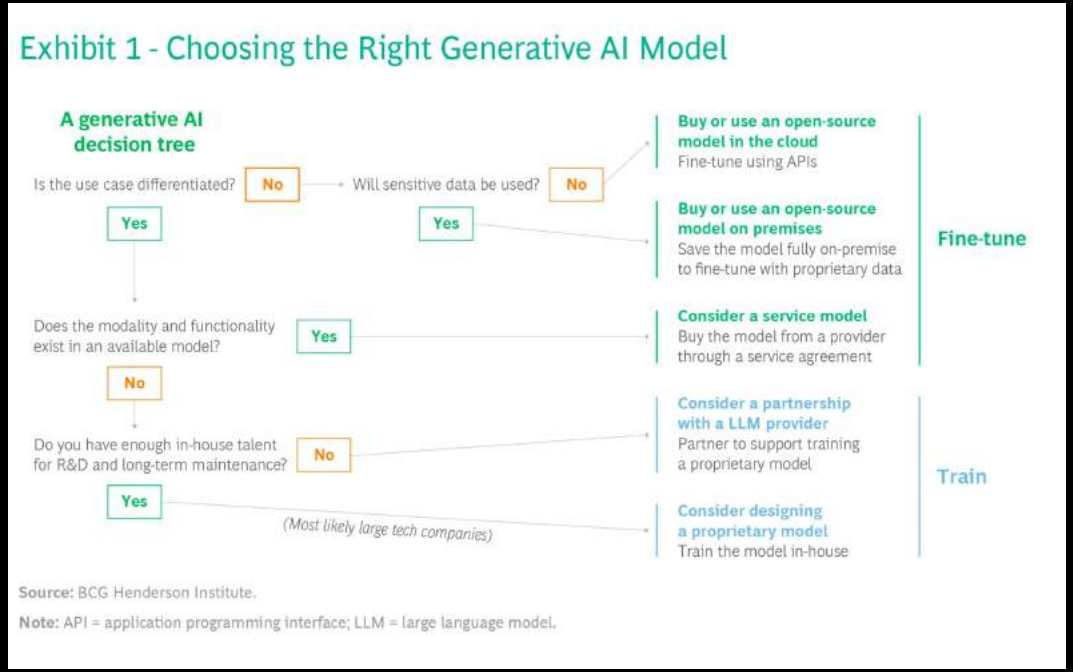
https://github.com/google/BIG-bench/
https://arxiv.org/pdf/2210.09261.pdf

NLP Existing Use Cases

Differentiated?

Data Sensitivity?

Common Use Case?



https://www.bcg.com/publications/2023/ceo-guide-to-ai-revolution

Longer impact is going to broader

- These LLMs can be the gateway to lots of other tools / services
- It's a way to 10X the value of your existing services by widening and easing access to them