

Tópicos Relevantes sobre o artigo "On the Opportunities and Risks of Foundation Models"

Nome: Caio Petrucci dos Santos Rosa

RA: 248245

Os principais tópicos que achei relevantes, durante a leitura, foram os seguintes:

- Os conceitos de *emergence* e *homogenization*:
 - *Emergence*: é quando o comportamento de um sistema é construído a partir de uma indução implícita em vez de uma construção explícita;
 - *Homogenization*: é quando uma metodologia específica de construção de sistemas aprendizado de máquina passa a ser aplicável e consolidada em uma vasta variedade de aplicações.
- Como, na história da pesquisa em IA, *emergence* e *homogenization* ocorreram:
 - **Machine Learning**:
 - *Emergence*: sistemas passam a ter a capacidade de descobrir **como** resolver uma tarefa a partir de dados;
 - *Homogenization*: diversas aplicações passaram a poder se basear em um algoritmo genérico de aprendizado com base em dados.
 - **Deep Learning**:
 - *Emergence*: tornou-se possível a aprendizagem de *higher-level features* através do processo de aprendizagem sem a especificação explícita dessas *features*;
 - *Homogenization*: as mesmas arquiteturas de modelos (como CNNs) passaram a ser aplicáveis para diversos tipos de aplicações.
 - **Foundation Models**:
 - *Emergence*: possuem níveis surpreendentes de *emergence* que surgiram conforme a escala desses modelos foi aumentando, o que levou a características como *in-context learning*;
 - *Homogenization*: possuem uma capacidade de homogeneização sem precedentes, pois se mostraram aplicáveis em diversas áreas sem modificação dos modelos.
- A definição sobre o que são **Foundation Models**:
 - São modelos que possuem gigantesco impacto social e econômico;
 - Servem como **fundação** para outros modelos e aplicações pois, apesar de serem incompletos por si só, servem como base para modelos *task-specific* através da alta capacidade de adaptação que **Foundation Models** possuem.
- Os riscos éticos de *deployment* de **Foundation Models** sem uma pesquisa profunda sobre as implicações éticas e seus impactos em aplicações diversas.
- A crítica dos autores sobre a relação entre o desenvolvimento tecnológico e as motivações da Indústria, assim como o apontamento pela necessidade de maior participação das universidades e de pesquisadores de diversas áreas, na proposição, desenvolvimento e pesquisa sobre novos modelos.

Assim, desde a escrita do artigo em julho de 2022, já ocorreram diversas mudanças como:

- Evolução do estado-da-arte: surgimento de diversos LLMs (ou **Foundation Models**), como o GPT-3.5, GPT-4, LLaMA, LLaMA 2, ...;
- Lançamento de diversas aplicações novas que tiram proveito desses **Foundation Models** através de *homogenization*, como ChatGPT, Midjourney e outras;
- Cada vez maior impacto social e econômico, dado diversas formas de trabalho já foram impactadas pelo surgimento dessas ferramentas: aumento de produtividade; diminuição de visibilidade de artistas (cujo os dados podem ter sido utilizados sem autorização); dentre outros impactos.