

## Relatório 5 - Leitura: Construindo uma LLM (II)

Vitor Eduardo de Lima Kenor

### Descrição da atividade

Nesse Card analisamos novamente o artigo “A Survey of Large Language Models”, desta vez foram analisados os tópicos 3 e 4. Iniciando a leitura pelo tópico 3, nele é abordado sobre ferramentas disponíveis para quem quer usar ou desenvolver modelos de LLM. Para quem quer criar soluções só que não quer criar um modelo próprio, tem as opções de modelos prontos que já são treinados e podem ser usados diretamente. Alguns dos modelos prontos são de código aberto como o LLaMa desenvolvido pela META, já outros são acessíveis por API como o GPT-4 da OpenAI. Modelos que podem ser acessados por API te dão a vantagem de não precisar rodar localmente, e modelos grandes necessitam de uma máquina muito potente. Mais pra frente no tópico 3, temos uma ideia de como os dados que formam a base de dados que treinam os modelos de LLM são escolhidos. Finalizando este tópico é citado ferramentas que facilitam o uso e o treinamento desses modelos sem precisar de conhecimento avançado em programação. Algumas das ferramentas que eu uso e me auxiliam muito hoje em dia são: Hugging Face, LangChain e ChatGrok.

Partindo para o tópico 4, temos uma explicação de como os modelos de LLM aprendem a gerar textos. Um dos principais fatores são os dados fornecidos aos modelos, por isso na etapa de coletar os dados eles precisam passar por alguns processos de limpeza e organização como é mostrado no artigo.

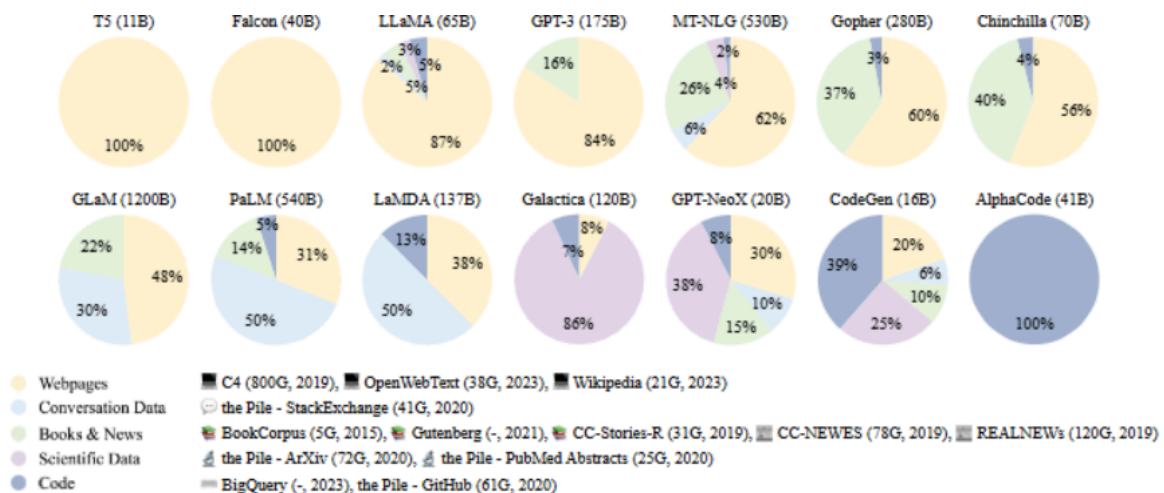


Fig. 6: Ratios of various data sources in the pre-training data for existing LLMs.

Para que o modelo entenda diferentes tipos de texto como notícias, diálogos e código de programação, ele é treinado com conteúdos variados. É explicado que no começo do treinamento o modelo treina com textos mais simples e depois passa para conteúdos difíceis, assim como nosso processo de aprendizagem. Por fim, um cuidado que se deve ter é a questão dos modelos criarem respostas muito repetitivas ou sem sentido, para resolver isso é feito o uso de técnicas que ajudam os modelos a equilibrar criatividade e coerência.

## **Conclusões**

Analizamos os tópicos do artigo onde o primeiro aborda ferramentas disponíveis para o uso e desenvolvimento de LLMs, falou sobre modelos prontos como o LLaMa e o GPT-4, para finalizar o tópico eu cito algumas ferramentas que me auxiliam no desenvolvimento de aplicações com LLM como Hugging Face, LangChain e ChatGroq. O outro tópico analisado, explica a forma como os LLMs aprendem a gerar texto, desde a coleta e organização dos dados até o treinamento progressivo, passando de textos simples para mais complexos.

## **Referências**

Link do artigo "A Survey of Large Language Models":

<https://arxiv.org/pdf/2303.18223>