

Questão 1

Máquina de aprendizagem Transformer

O que é modelo Transformer

O modelo transformer é uma rede neural que é capaz de aprender sobre o contexto como também o significado com o monitoramento de relações em dados sequenciais. Esse tipo de modelo é baseado na autoatenção, ou seja, ele é capaz de detectar de como os elementos de dados distantes podem estar relacionados uns aos outros.

Os transformers funcionam da seguinte forma: Dados como texto, imagens, sinais e falas são recebidos por eles, em seguida, essas redes neurais passam pelo processo de treinamento e aprendizagem através dos dados recebidos e, no final, passam pelo processo de adaptação com a intenção de realizar seguintes tarefas como responder questões recebidas por meio de texto (um exemplo é perguntar ao ChatGPT sobre um determinado assunto e essa ferramenta de Inteligência Artificial trazer uma resposta sobre essa pergunta), analisar sentimentos, extrair informações, capturar imagens, reconhecer objetos ou, até mesmo, seguir instruções. Um exemplo de procedimento feito para os transformers é apresentado na figura 1.

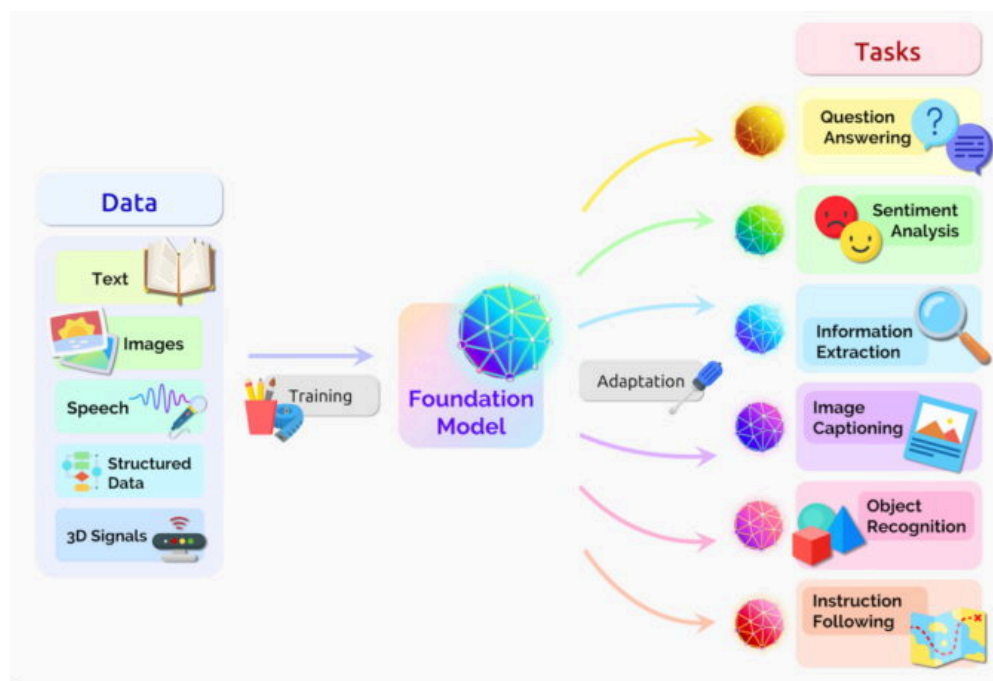


Figura 1: Esquemático apresentando o processo de treinamento dos transformers

Fonte: Nvidia

Atualmente esse tipo de máquina de aprendizagem está bastante abrangente no dia a dia. Pessoas estão cada vez mais usando esse tipo de modelo para fins

escolares, entretenimento, tirar dúvidas ou, até mesmo, para realizar pesquisas de saúde ou detectar fraudes. Existem várias ferramentas de aprendizagem que utilizam os modelos transformers como ChatGPT, Meta AI, Copilot, Google Gemini e entre outros.

Existem várias aplicações sobre a máquina de aprendizagem transformer. Duas dessas aplicações serão abordadas nos tópicos a seguir.

Conversão de texto para voz (text to speech)

Uma das aplicações bastante comuns no processo de aprendizagem transformer é a conversão de texto digitado para voz. Nela o usuário digita qualquer texto e clica para que uma determinada voz seja gerada seguindo a mensagem de texto digitado. Existem várias ferramentas que realizam esse tipo de conversão de texto para voz, como o Falatron, o TTSMaker, ElevenLabs, entre outros.

Alguns exemplos de ferramenta de conversão de texto para voz estão disponíveis nestes sites: <https://ttsmaker.com/br> e <https://speechgen.io/pt/>. Algumas ferramentas possuem opção de personalização de voz gerada, como voz masculina, voz feminina, tipo de sotaque, idioma, tom e timbre.

Geração de Imagens

Outra aplicação que está sendo muito utilizada, principalmente nas redes sociais, é o uso da ferramenta de transformer para geração de imagens. Dentro do aplicativo ou site que realiza esse procedimento, basta digitar na caixa de texto a descrição de como deve ser a imagem (instruções) para que a ferramenta realize a geração. O ChatGPT utiliza como modelo transformer para geração de imagens o DALL-E 3.

Sites pesquisados:

<https://blog.nvidia.com.br/blog/o-que-e-um-modelo-transformer/>