

Azure OpenAI por Renan Ribeiro

Entendimento de como compreender e utilizar os serviços de IA da Azure com base na primeira etapa do bootcamp DIO

1. Introdução ao Azure OpenAI

O Azure Open AI é uma plataforma avançada que oferece uma ampla gama de serviços de Inteligência Artificial (IA) para desenvolvedores e organizações. Esta plataforma permite a implementação rápida e eficiente de modelos de IA sofisticados, proporcionando soluções tecnológicas inovadoras.

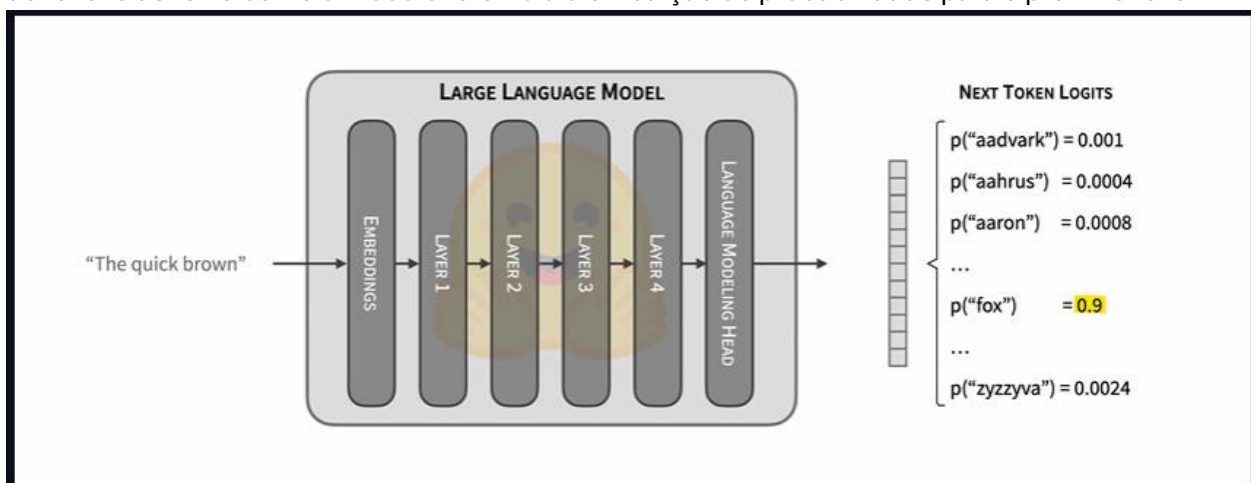
Dentro da plataforma é possível criar instâncias para serem usadas de forma independente (Seja via API ou Playground) dentro do Azure AI Foundry, possibilitando seguir custos relacionados a inputs e outputs de tokens dentro da plataforma em seus modelos.

2. Modelos de Linguagem de Grande Escala (LLM)

Para um bom entendimento da capacidade e potencial dos modelos criados dentro da plataforma, é importante entender os Modelos de Linguagem de Grande Escala (LLM) que constituem um componente essencial do Azure OpenAI. Desenvolvidos para processar vastas quantidades de dados textuais, esses modelos geram respostas inteligentes e contextualmente relevantes baseado em “tokens prováveis”. Utilizando técnicas avançadas de aprendizado de máquina, os LLMs compreendem e geram linguagem natural de maneira eficaz.

Exemplos:

1. Um modelo de linguagem treinado para modelagem de linguagem causal recebe uma sequência de tokens de texto como entrada e retorna a distribuição de probabilidade para o próximo token.



2. Em uma estrutura de LLM, a janela de contexto de informações acabam sendo maiores que em um Modelo de Linguagem de Baixa Escala (SLM) possibilitando uma maior capacidade de geração de respostas mais elaboradas. Em um cenário onde no modelo SLM é possível conseguir

informações mais diretas como “Clima e Tempo atual”, em uma LLM a resposta poderia vir com alguma sugestão de lugares a se visitar/evitar de acordo com as informações fornecidas.

3. Exemplos de Tokenização

Dentro do LLM, a tokenização é o processo de segmentação do texto em unidades menores chamadas tokens, que podem ser palavras, caracteres ou subpalavras. Essa etapa é crucial no processamento de linguagem natural, permitindo que os modelos de IA compreendam e manipulem o texto de forma granular.

É possível entender um pouco melhor como os tokens são gerados com base nos textos, na plataforma Tokenizer da OpenAI. O uso de idiomas diferentes do nativo do modelo (EN), pode gerar em uma perda de performance aumentando a quantidade de tokens gerados para frases dentro do mesmo contexto.

- Exemplo de tokenização 1 (EN):

Tokens	Characters
3	15
Do an elevation	

- Exemplo de tokenização 2 (PT):

Tokens	Characters
4	17
Faça uma elevação	

4. Modelos de IA fornecidos pelo Azure

O serviço Azure Open AI oferece uma diversidade de modelos de IA adequados a diferentes finalidades. Alguns dos modelos mais destacados incluem:

- GPT-4o mini: Um modelo de linguagem natural avançado capaz de realizar uma ampla gama de tarefas, desde a geração de texto até a tradução de idiomas. Exemplo: Traduzir em um chat mensagens recebidas em um idioma para outro de forma instantânea.
- Codex: Um modelo especializado em programação, projetado para gerar código e auxiliar desenvolvedores em tarefas de codificação. Exemplo: Poderia ser usado desde suporte em desenvolvimento até na rescrita de códigos em outras linguagens.
- DALL-E: Um modelo de geração de imagens que cria ilustrações realistas a partir de descrições textuais. Exemplo: Modelo pode ser utilizado para prototipação de ideias visuais como, Cenários, Logotipos, Mascotes, Banners e etc.
- CLIP: Um modelo que compreende e associa texto e imagens, possibilitando pesquisas multimodais e outras aplicações inovadoras. Exemplo: Extrair textos de imagens podendo

ser úteis em tarefas como: Leitura de documentos para que outra IA possa resumir ou até inserir em uma base de conhecimento, validação de dados de documentos, leitura de placas, identificação/descrição de objetos na imagem e etc.

5. Conclusão

O Azure Open AI se destaca como uma plataforma robusta e versátil, oferecendo uma ampla gama de modelos de Inteligência Artificial que atendem a diversas necessidades e aplicações. Desde a geração de texto com o GPT-4o mini até a criação de imagens com o DALL-E, a plataforma possibilita a implementação de soluções inovadoras e eficientes. A compreensão e utilização dos Modelos de Linguagem de Grande Escala (LLM) são fundamentais para aproveitar ao máximo as capacidades do Azure OpenAI, permitindo a geração de respostas contextualmente relevantes e a manipulação granular de texto através da tokenização. Com esses recursos, desenvolvedores e organizações podem explorar novas fronteiras tecnológicas e otimizar processos, tornando-se mais competitivos no mercado.