**Infnet – Ciência de Dados – 2024.2**

Engenharia de Prompts para Ciência de Dados

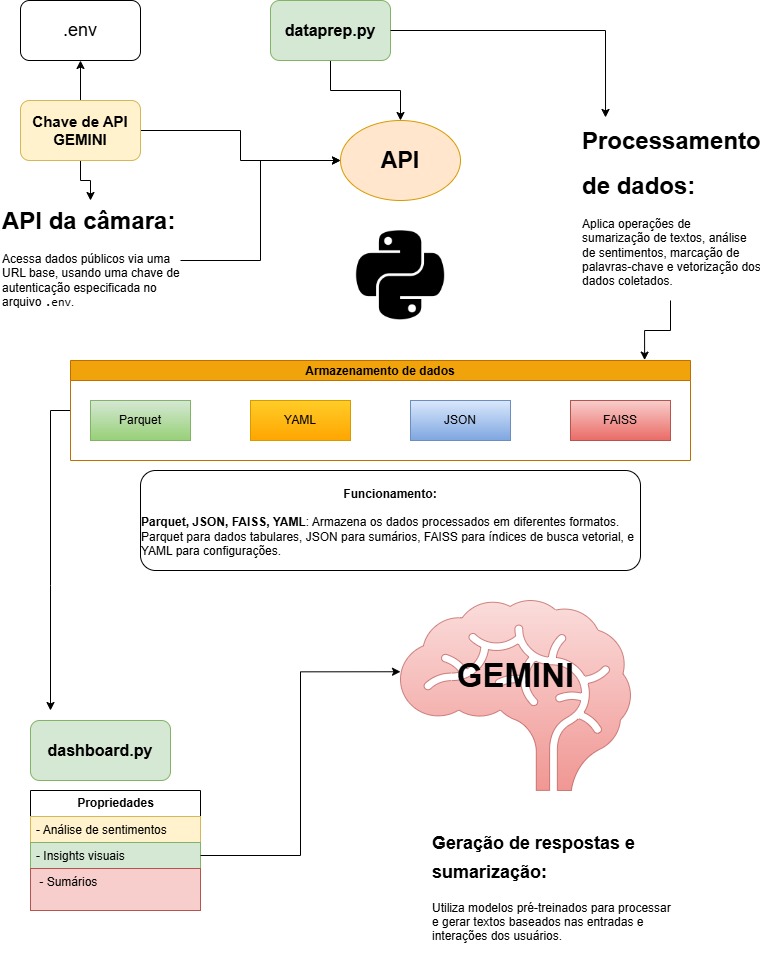
AT – Questões Escritas

Prof.: Thiago Ciodaro Xavier

Aluno: Daniel da Cunha Moreira

**Questão 1:**

Letra a)



Letra b)

**Arquitetura do Projeto de Processamento de Dados da Câmara dos Deputados**

1. **Coleta de Dados (dataprep.py)**

* **API da Câmara dos Deputados**: Utilizada para extrair dados estruturados sobre atividades legislativas, como proposições e despesas dos deputados.
* **Arquivo .env para chave Gemini**: Armazena chaves de configuração e API de forma segura, evitando exposição no código fonte.
* **Processadores de dados**:
  + **Sumarização**: Utiliza LLMs para condensar informações extensas em resumos concisos.
  + **Análise de sentimentos**: Avalia o tom emocional (positivo, negativo, neutro) das declarações ou documentos.
  + **Marcação de palavras-chave**: Identifica e marca termos e conceitos relevantes nos textos.
  + **Vetorização para bases de referência**: Transforma textos em vetores numéricos que representam semanticamente o conteúdo para facilitar comparações e recuperações rápidas.

**2. Armazenamento de Dados**

* **Parquet para dados tabulares**: Formato eficiente para armazenamento e consulta de dados estruturados em colunas.
* **JSON para sumários e metadados**: Utilizado para armazenar resumos e informações adicionais em formato de fácil manipulação e acesso.
* **FAISS Index para buscas vetoriais**: Permite a realização de buscas de alta performance em grandes volumes de dados vetorizados.
* **YAML para configurações**: Armazena configurações de forma estruturada, facilitando a gestão e interpretação por humanos e máquinas.

**3. Dashboard (dashboard.py)**

* **Abas de Visualização**: Oferece interfaces separadas para visualizar resumos, análises de sentimentos e resultados de pesquisas vetoriais.
* **Funcionalidades interativas e uso de LLMs**: Permite que os usuários interajam com os dados e solicitem processamentos ad hoc, como sumarizações e análises, utilizando LLMs integrados.

**4. Modelos de Linguagem de Larga Escala (LLMs)**

* **Sumarização e geração de respostas**: Implementados para processar e responder a consultas complexas, apoiando decisões e análises legislativas. Utilizam a arquitetura de **Transformers**, que são modelos de atenção que aprendem contextos em larga escala de dados textuais. Os Transformers são fundamentais para entender a estrutura linguística e gerar respostas contextuais precisas.

Letra c)

**Funcionamento de LLMs e Aplicações em Sumarização**

**LLMs (Large Language Models)**, como os baseados na arquitetura Transformer, são modelos de inteligência artificial treinados para entender e gerar linguagem humana. Eles operam através de um mecanismo de atenção que permite ao modelo considerar o contexto de cada palavra no texto, independentemente da distância entre as palavras no texto. Isso é crucial para entender a estrutura gramatical e o significado semântico.

Na sumarização, os LLMs podem ser empregados de duas formas principais:

1. **Sumarização Extrativa:** Identifica as partes mais importantes do texto (frases ou segmentos) e as compila para formar um resumo. Esses modelos dependem fortemente da capacidade do modelo de avaliar a relevância e a representatividade de segmentos de texto dentro de um documento maior.
2. **Sumarização Abstrativa:** Gera um novo texto que captura as ideias principais do conteúdo original, frequentemente reescrevendo ou parafraseando para produzir um resumo mais conciso e coeso. Isso envolve um entendimento mais profundo e a capacidade de reformular informações, algo para o qual os LLMs são particularmente bem adaptados devido à sua capacidade de geração de linguagem.

**Questão 2:**

**Prompt Utilizado em todos os modelos:**

“Gere um texto curto de dois parágrafos que explique a câmara dos deputados.”

**Respostas dos 3 LLMs utilizados:**

**Claude-3.5-Sonnet**

A Câmara dos Deputados é uma das duas casas do Congresso Nacional brasileiro, sendo composta por 513 deputados federais eleitos pelo povo para mandatos de quatro anos. Estes parlamentares representam os diferentes estados brasileiros e o Distrito Federal, com o número de representantes de cada unidade federativa sendo proporcional à sua população.

Como parte do Poder Legislativo, a Câmara tem como principais atribuições criar e aprovar leis, fiscalizar o Poder Executivo e debater questões de interesse nacional. Os deputados também são responsáveis por analisar o orçamento da União, propor emendas à Constituição e julgar crimes de responsabilidade cometidos por autoridades, como o Presidente da República, em conjunto com o Senado Federal.

**Gemini-1.5-FLash**

A Câmara dos Deputados é uma das duas casas do Congresso Nacional brasileiro, sendo a outra o Senado Federal.  Ela é composta por deputados federais, eleitos por representação proporcional em cada estado, com o número de deputados por estado sendo proporcional à sua população.  A principal função da Câmara é legislar, ou seja, criar leis federais que regem o país.  Além disso, a Câmara também tem o poder de aprovar o orçamento anual do governo, fiscalizar as ações do Poder Executivo e julgar o Presidente da República em casos de crime de responsabilidade.

A representatividade da Câmara dos Deputados garante que diferentes regiões e grupos populacionais tenham voz no processo legislativo.  Os deputados são eleitos para mandatos de quatro anos, podendo ser reeleitos.  O funcionamento da Câmara é regido por um Regimento Interno, que define as normas de funcionamento e procedimentos para votações e debates.  A Câmara desempenha um papel fundamental na democracia brasileira, sendo um espaço de debate e construção de políticas públicas.

**GPT-4o-Mini**

A Câmara dos Deputados é uma das duas casas do Congresso Nacional do Brasil, composta por 513 deputados federais eleitos pelo povo para mandatos de quatro anos. Os deputados representam diferentes estados e o Distrito Federal, com o número de representantes de cada unidade federativa sendo proporcional à sua população. A principal função da Câmara é elaborar e aprovar leis que regem o país, além de fiscalizar as ações do Poder Executivo e debater questões de interesse nacional.

Além de legislar, a Câmara dos Deputados também é responsável por analisar o orçamento da União, propor emendas à Constituição e julgar crimes de responsabilidade cometidos por autoridades, como o Presidente da República. O funcionamento da Câmara é regido por um Regimento Interno, que define normas e procedimentos para votações e debates, garantindo que a diversidade de vozes e interesses da população seja representada no processo legislativo.