

Extractive Q&A

Performance Comparison between Learning Methods Context and Transfer Learning

Leonardo Augusto da Silva Pacheco
Marcus Vinícius Borela de Castro

Extractive Answering (The answer to a question given a piece of context is a direct substring of the context (fragmento do texto) - Analogy; A highlighter)

Sem considerar o retriever, que seria um "limitador" de performance

Abstractive Answering-The answer to a question given a piece of context is a free-form phrase based on the context-Analogy - A pen

Exemplo

[Inglês SQuAD 1.1](#)

[Português \(SQuAD_v1_1_pt\)](#)

[Stanford Question Answering Dataset](#) (base de teste)

A binary metric that gives $EM = 1$ if the characters in the predicted and ground truth answers match exactly, and $EM = 0$ otherwise. If no answer is expected, the model gets $EM = 0$ if it predicts any text at all.

Measures the harmonic mean of the precision and recall. (considerar palavras coincidentes e diferentes)

EM@3

F1@3

Model (261 mb) is a fine-tune checkpoint of DistilBERT-base-cased, fine-tuned using (a second step of) knowledge distillation on SQuAD v1.1 (F1=87.1 on the dev set) (for comparison, BERT bert-base-cased reaches F1= 88.7).

distilbert-base-cased-distilled-squad — Inglês

Transfer Learning

Modelo Q&A (1.33 gb)
com refinamento no
squad_v1_pt. (F1:84.4 ;
EM=72.68)

[pierreguillou/bert-large-cased-squad-v1.1-portuguese](#)

Português (baseados no BERTimbau)

Modelos

[GPT-neo_1.3B](#)

[GPT-J_6B](#)

[GPT-NEOx-20b](#)

Um desses: <https://huggingface.co/EleutherAI>

Context Learning

Resultados preliminares - verificação e tratamento dos dados

Totais artigos:48 paragrafos:2067 perguntas:10570 respostas:34726
Erros perguntas: repetidas 0
Erro respostas fora do texto: 0 fora da posição 0

Inglês

Erros perguntas: repetidas 2505

português

Avaliação da qualidade dos datasets

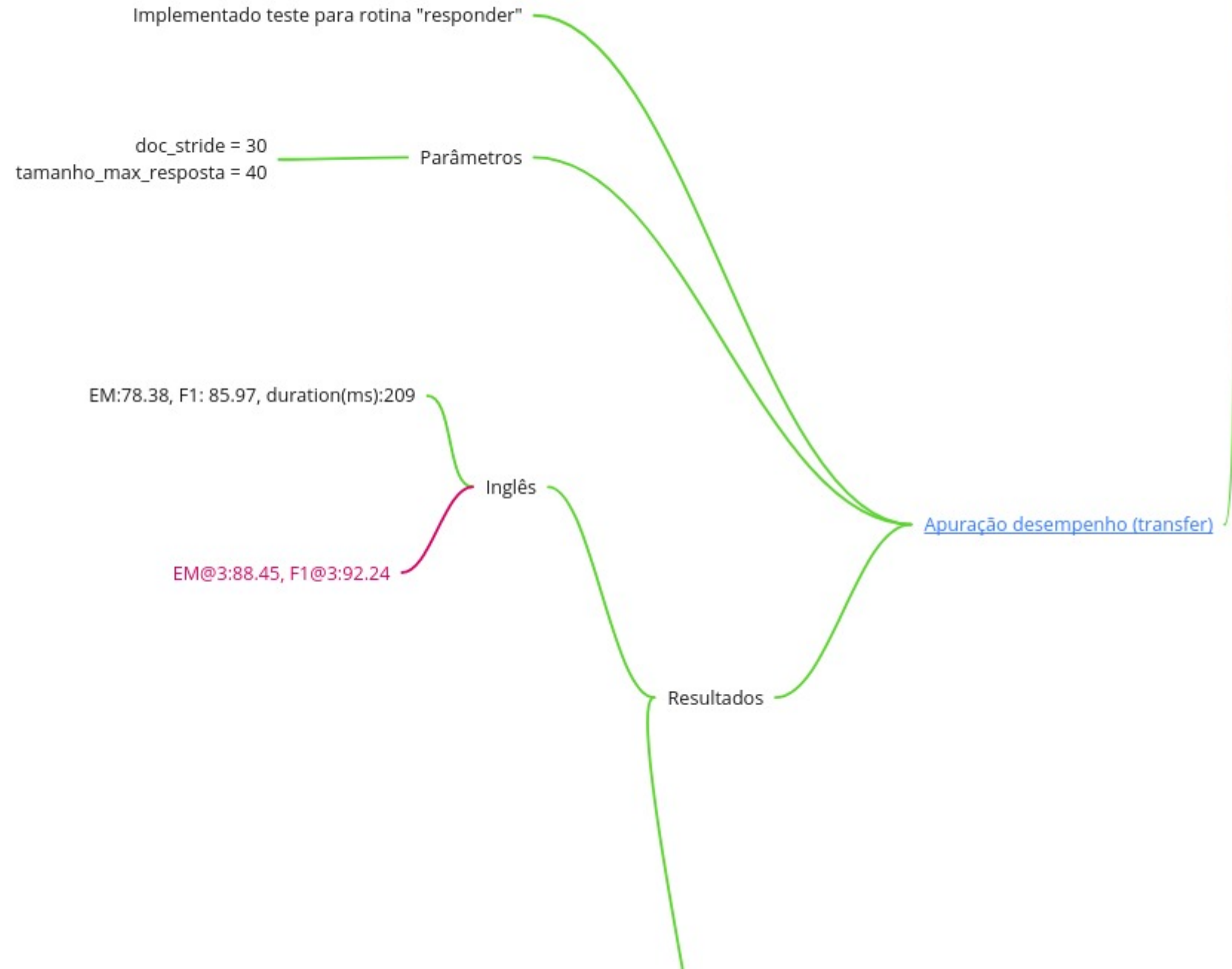
Datasets

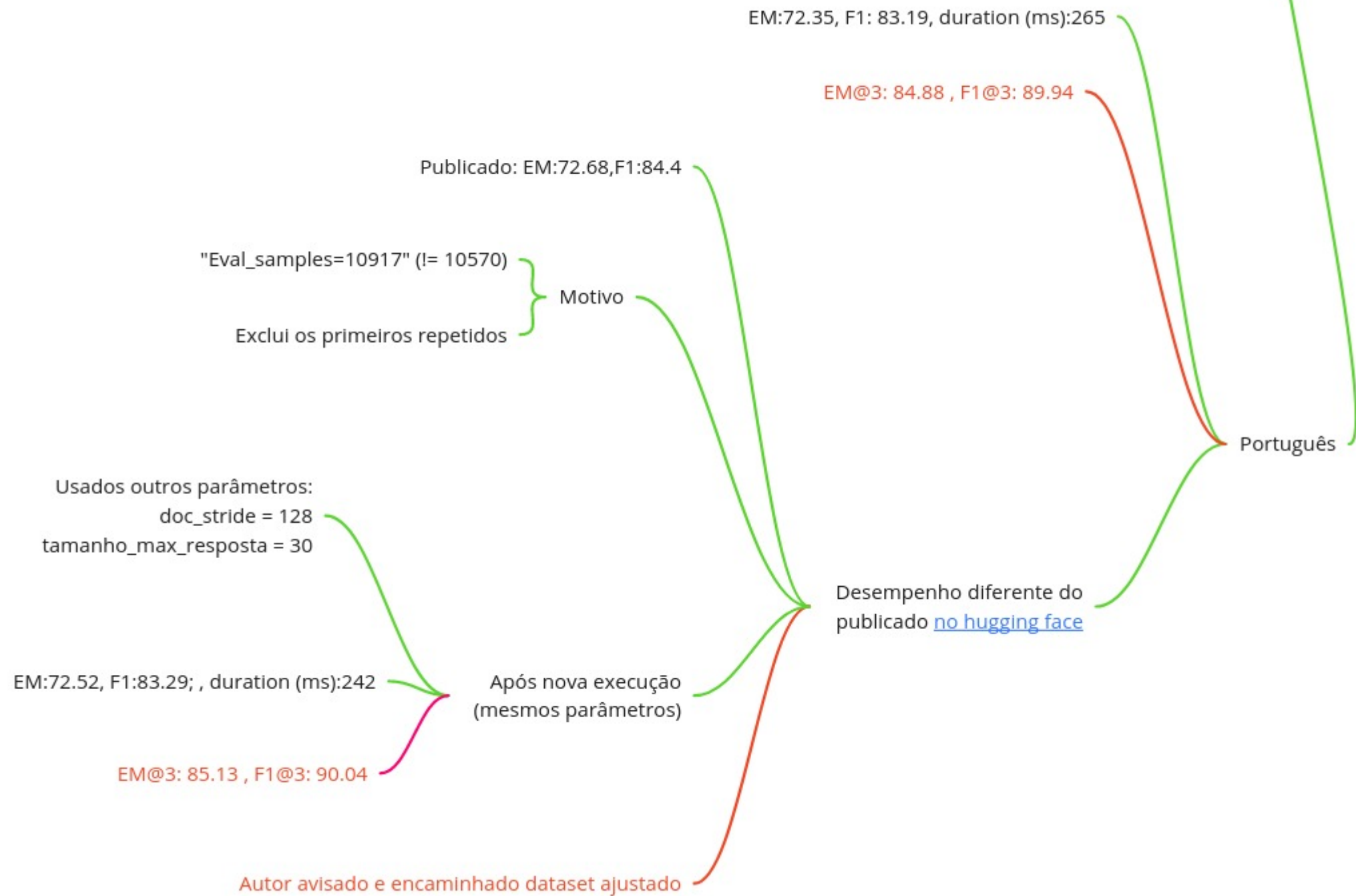
Dataset foi traduzido de forma automática e houve processo de revisão

Totais artigos:48 paragrafos:5644 perguntas:13075 respostas:34726
Erros perguntas: repetidas 2505 excluídas 2505

Exclusão de perguntas repetidas dataset português

Resultados preliminares - Transfer Learning





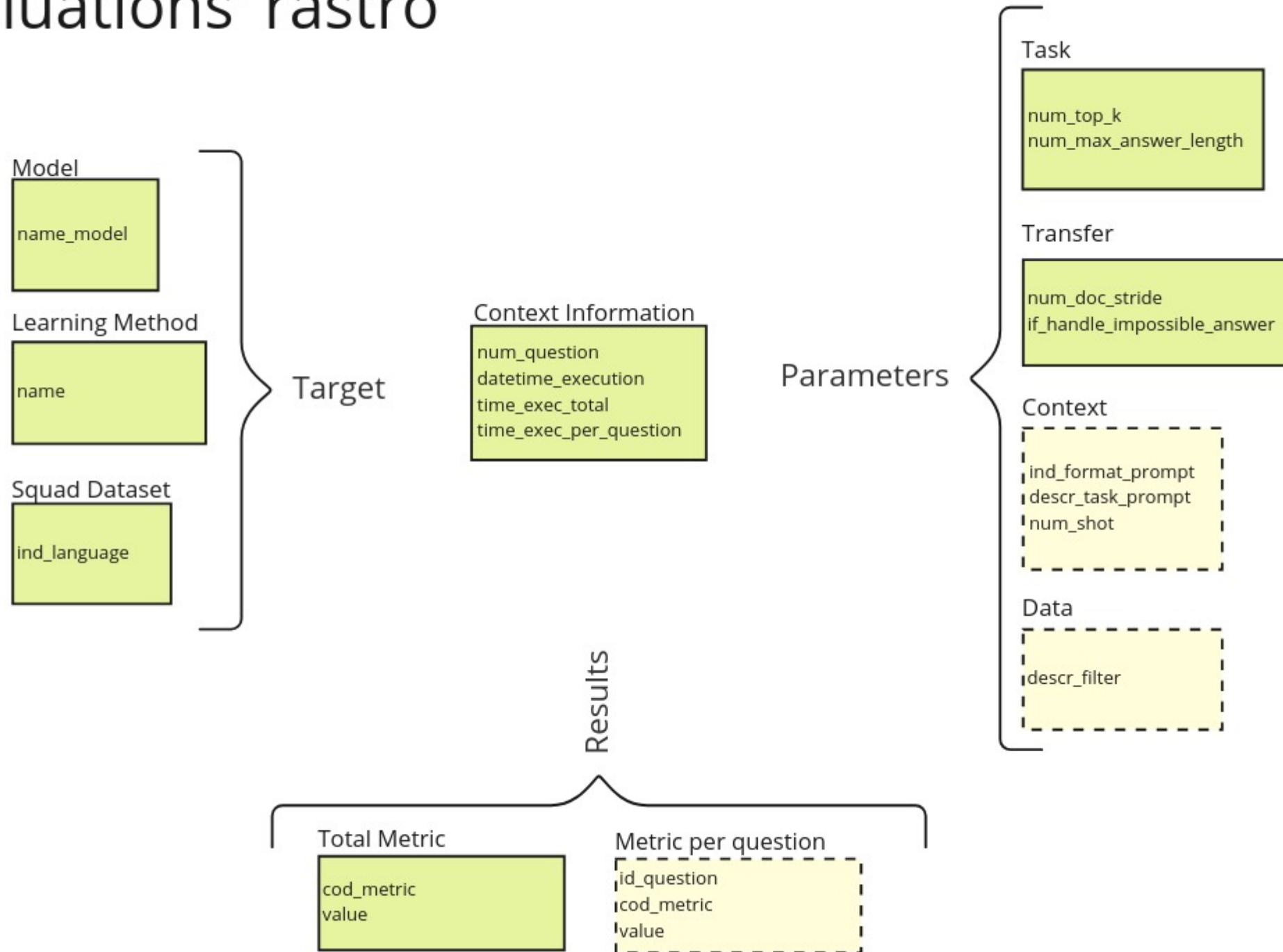
“Modelar um conhecimento colabora com o amadurecimento da aprendizagem e facilita a organização de processos e a estruturação de conceitos.”

Castro, Marcus 🤔👋



(imagem: brain-3829057__340_pixabay)

Evaluations' rastró*



* Pandas Dataframe

GPT-3

para definir formato prompt e parâmetros



Primeiras experimentações (Context)

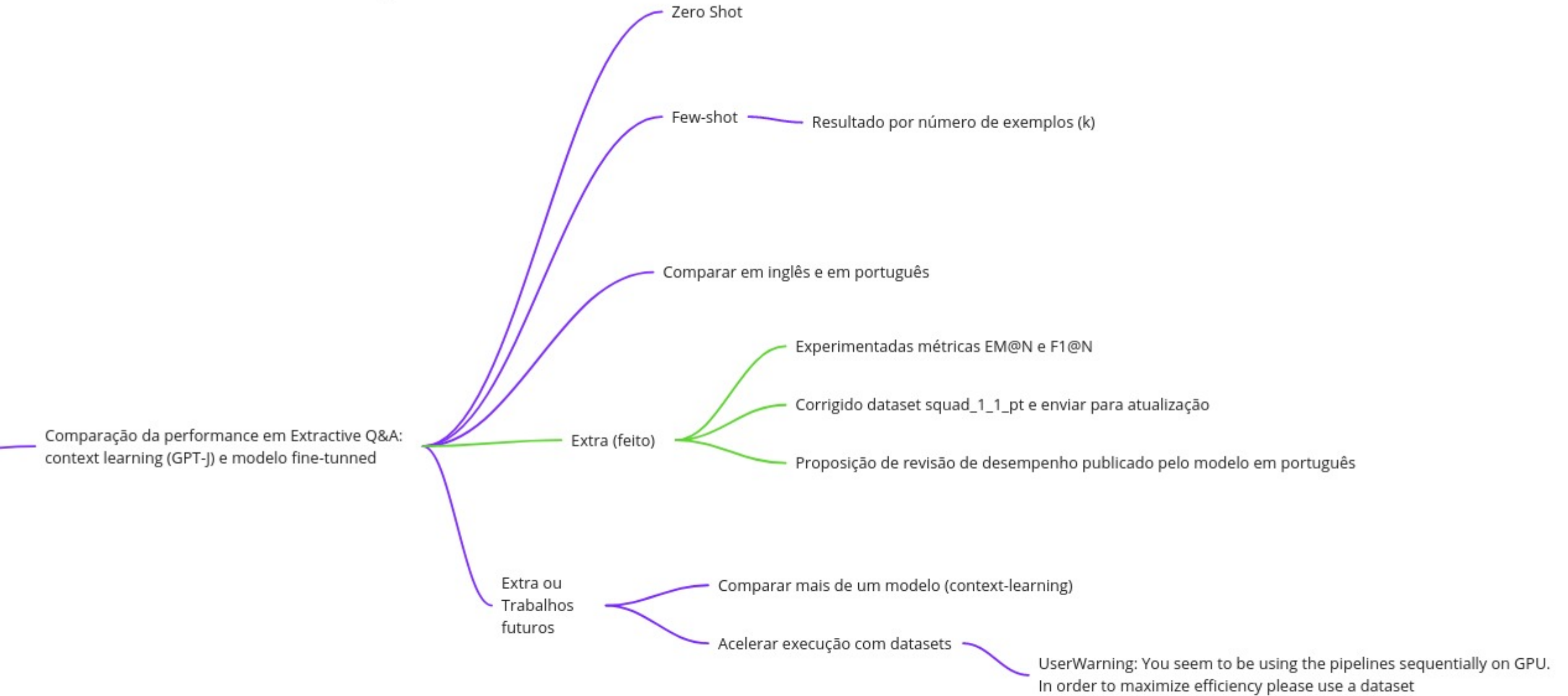
Instrução: Dado o contexto, dê a resposta à pergunta:\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: De onde os documentos foram retirados?\nResposta: Do tribunal\n\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: Quando tomei posse?\nResposta: 1990\n\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: Qual o melhor time do Brasil?\nResposta: Flamengo\n\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: Quando tomei posse?\nResposta: 1990\n\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: Quando tomei posse?\nResposta: 1990\n\nContexto: Por causa da chuva, o Flamengo, melhor time do Brasil, ficou sem jogar a final da Libertadores. Há muitos torcedores pelo Flamengo, o melhor time, na minha Família. Além da bicicleta, em 1990, eu tinha a honra de ser flamenguista. Ontem, quando reli os documentos do tribunal, descobri que em 1990 quando tomei posse, minha declaração de bens só continha uma bicicleta.\nPergunta: Quando tomei posse?\nResposta: 1990

[GPT-NEOx-20b](#)

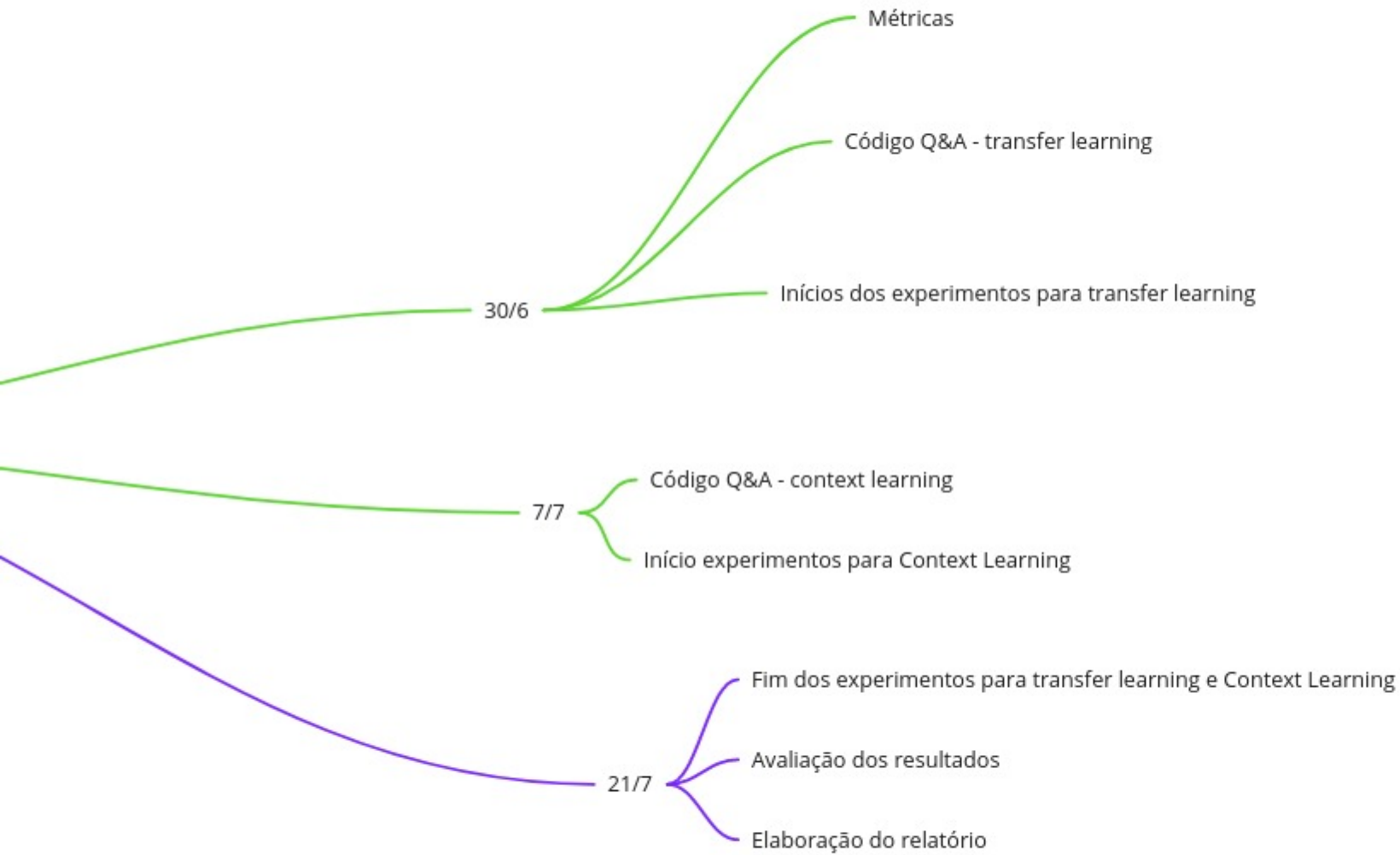
erro memória RAM



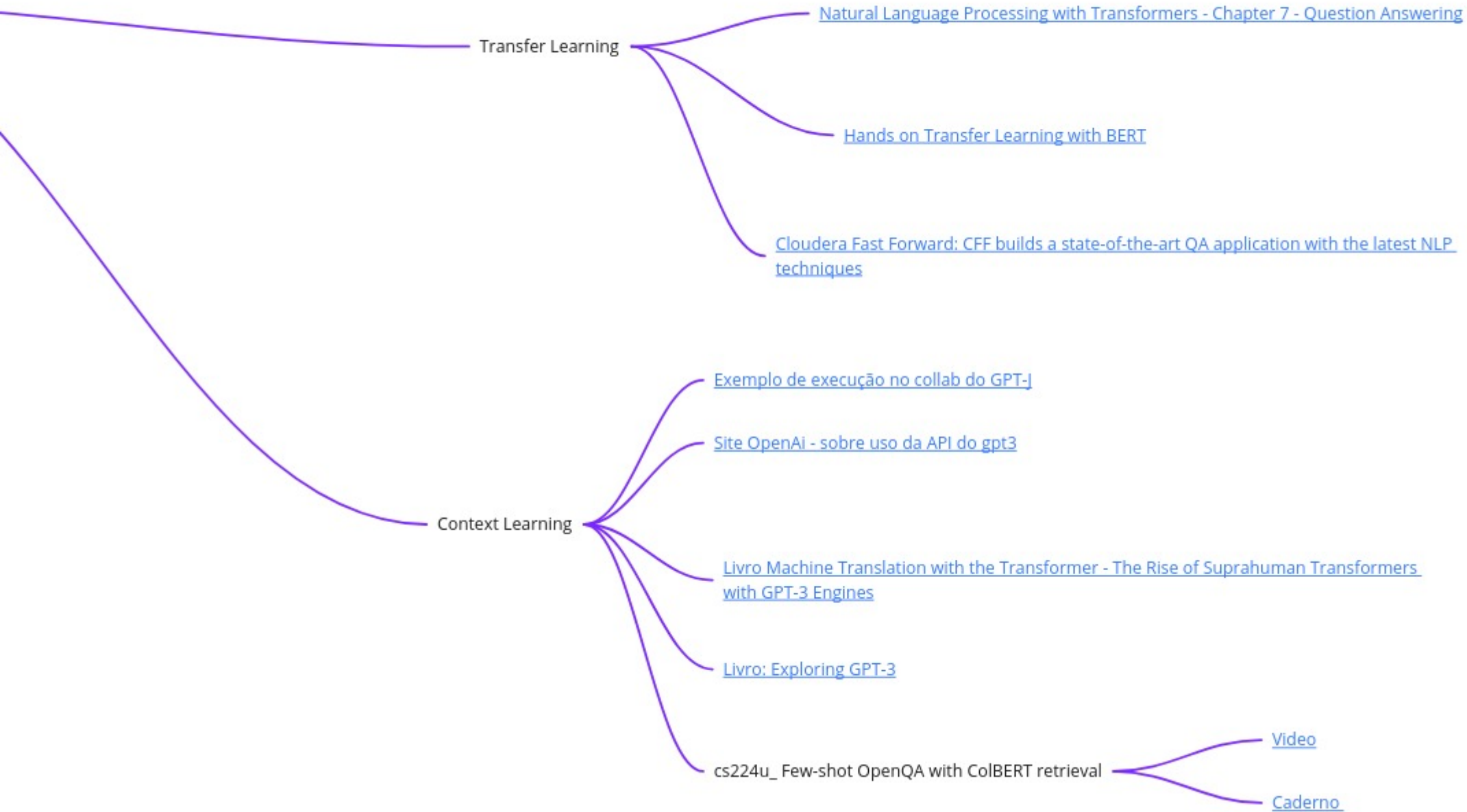
Resultados esperados



Cronograma

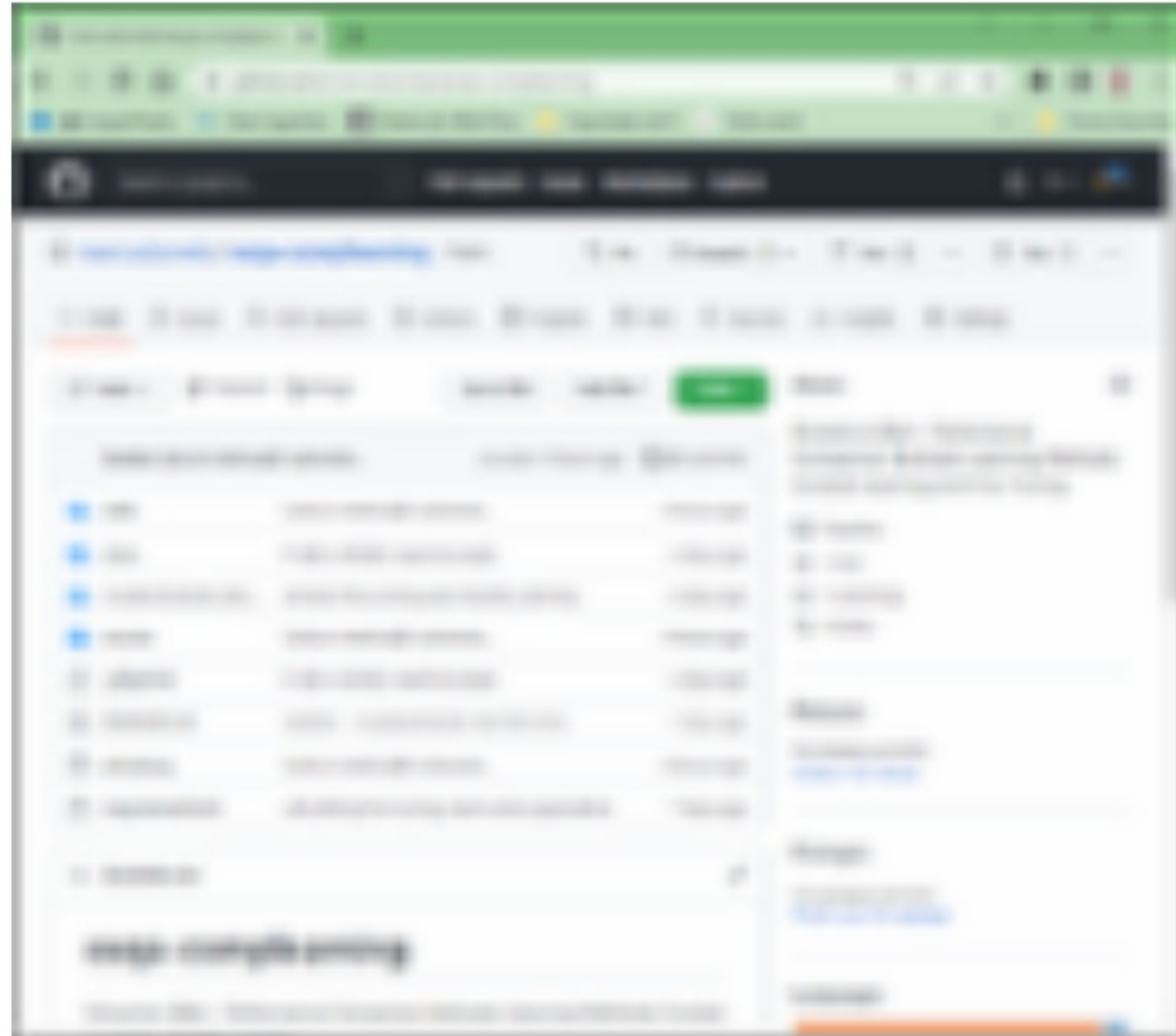


Referências



FIM

[Github do Projeto](#)



Não é contexto: abstractive Q&A

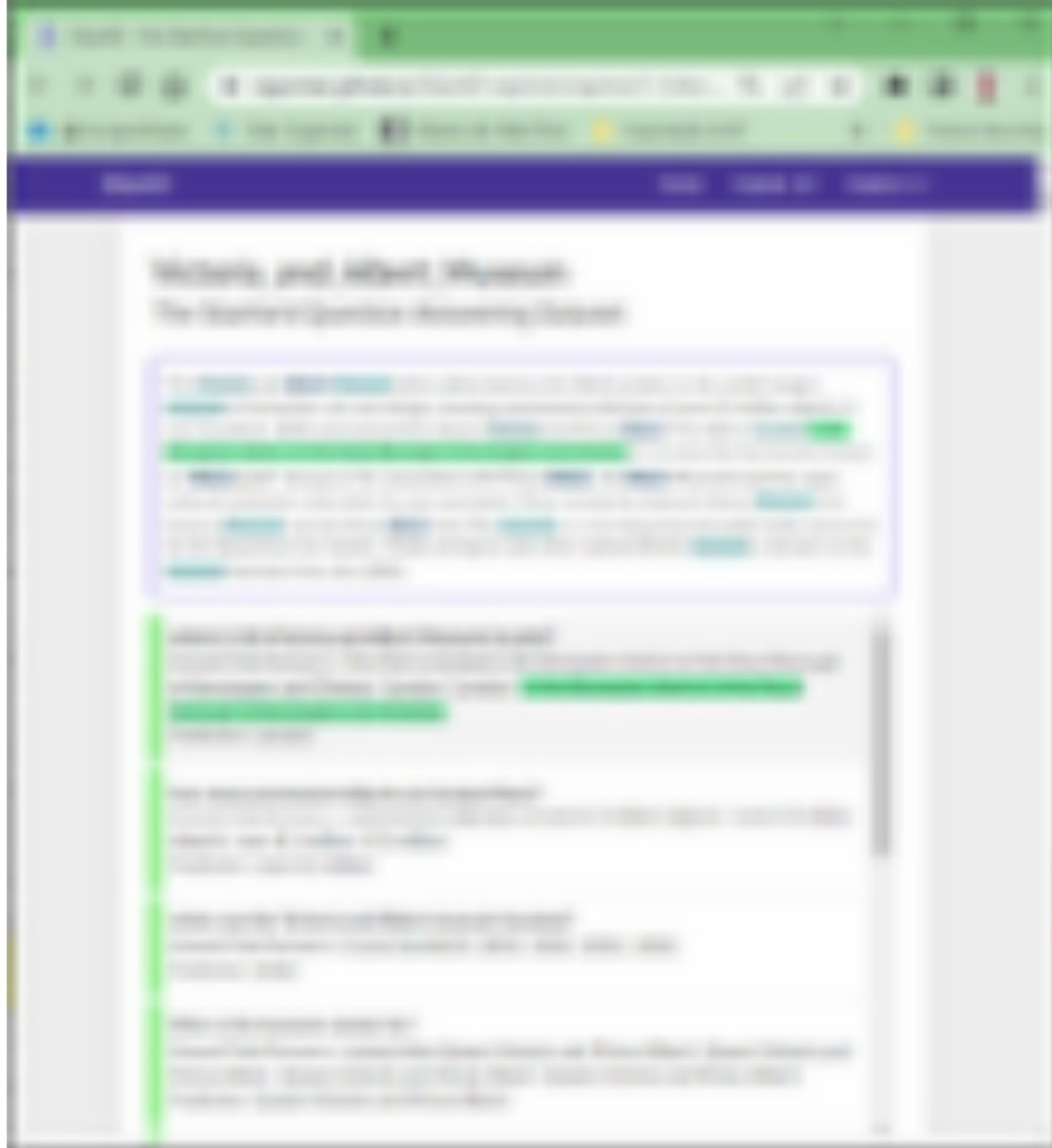


Extractive Q&A - Performance Comparison between
Learning Methods:
Context Learning and Transfer Learning

Resultados Desejados

Descrição

O que é Extractive Q&A?



Exemplo: [Stanford Question Answering Dataset SQuAD](https://quadrant.stanford.edu/)

IR Extractive Q&A (over Information Retrieval)



Fonte: [How to Build an Open-Domain Question Answering System?](#)

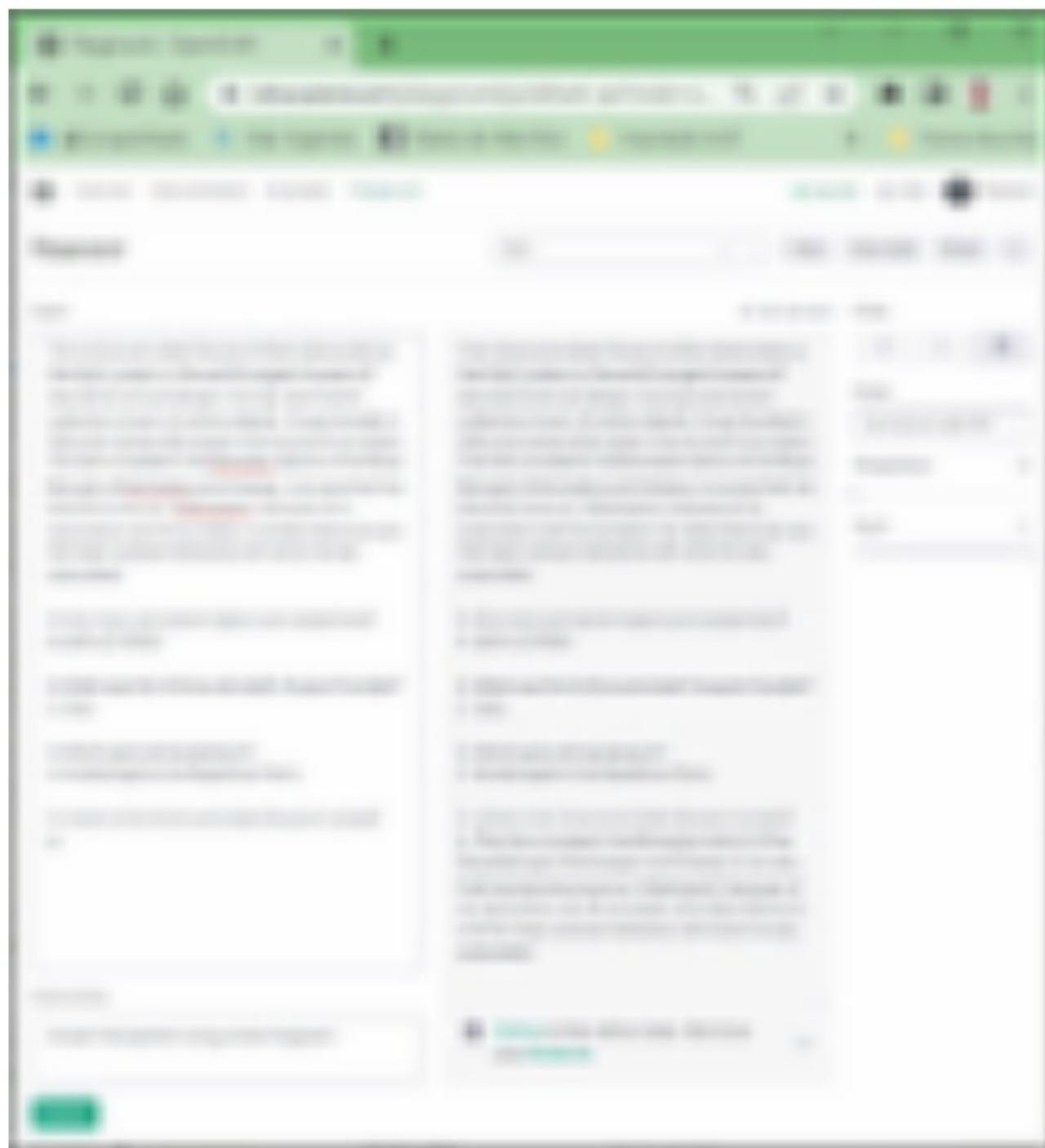
Extractive Q&A in Fine Tuning



Q&A in Context Learning

Contexto

Few Shot
k=3



Exemplo: [GPT-3](#) x SQuAD

Igualdade de condições: sem retrieval (sem limitador de performance)

