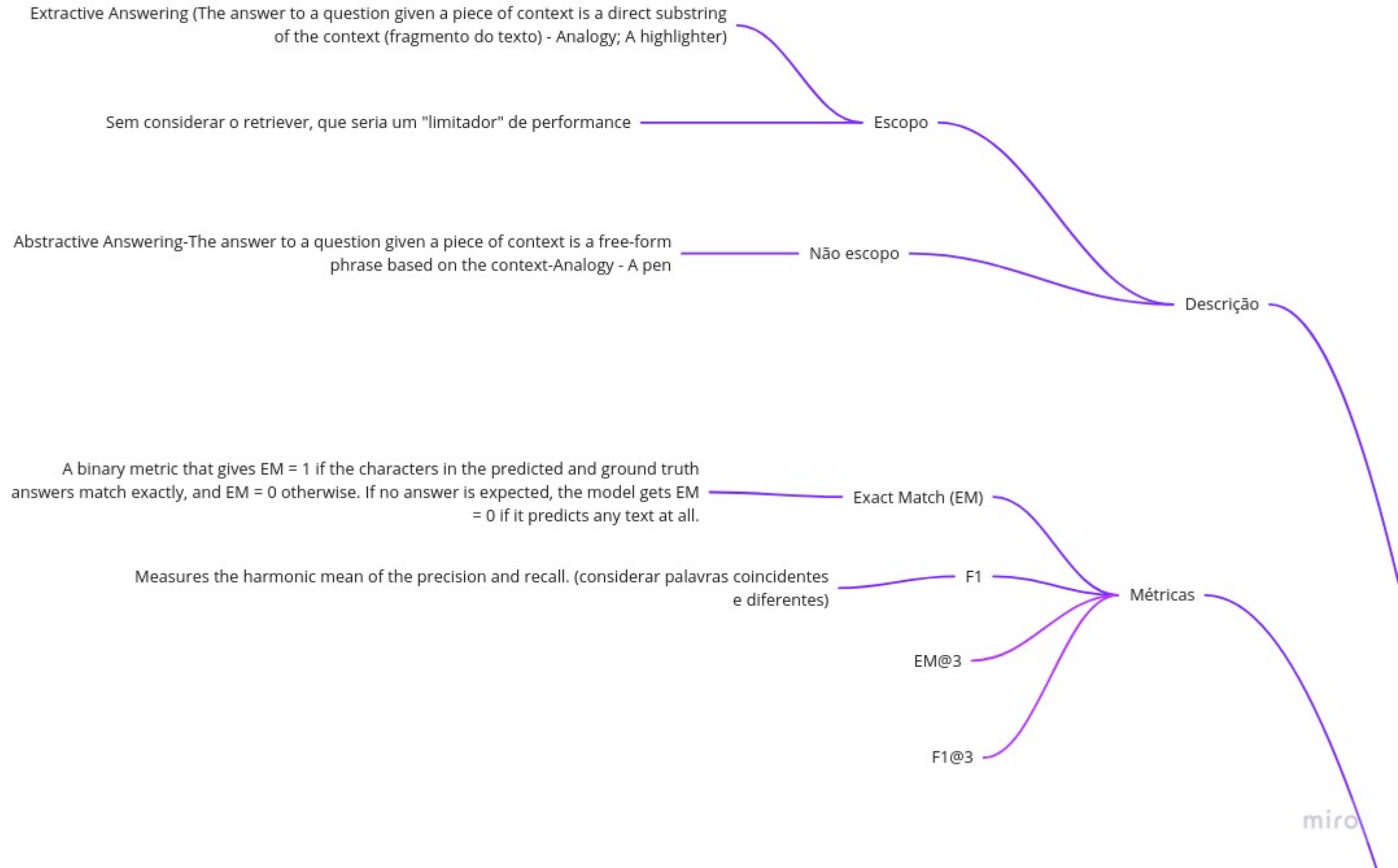


Extractive Q&A

Performance Comparison between Learning Methods Context and Transfer Learning

Leonardo Augusto da Silva Pacheco
Marcus Vinícius Borela de Castro



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

Modelos

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](https://huggingface.co/pierreguillou/bert-large-cased-squad-v1.1-portuguese)

Português (baseados no BERTimbau)

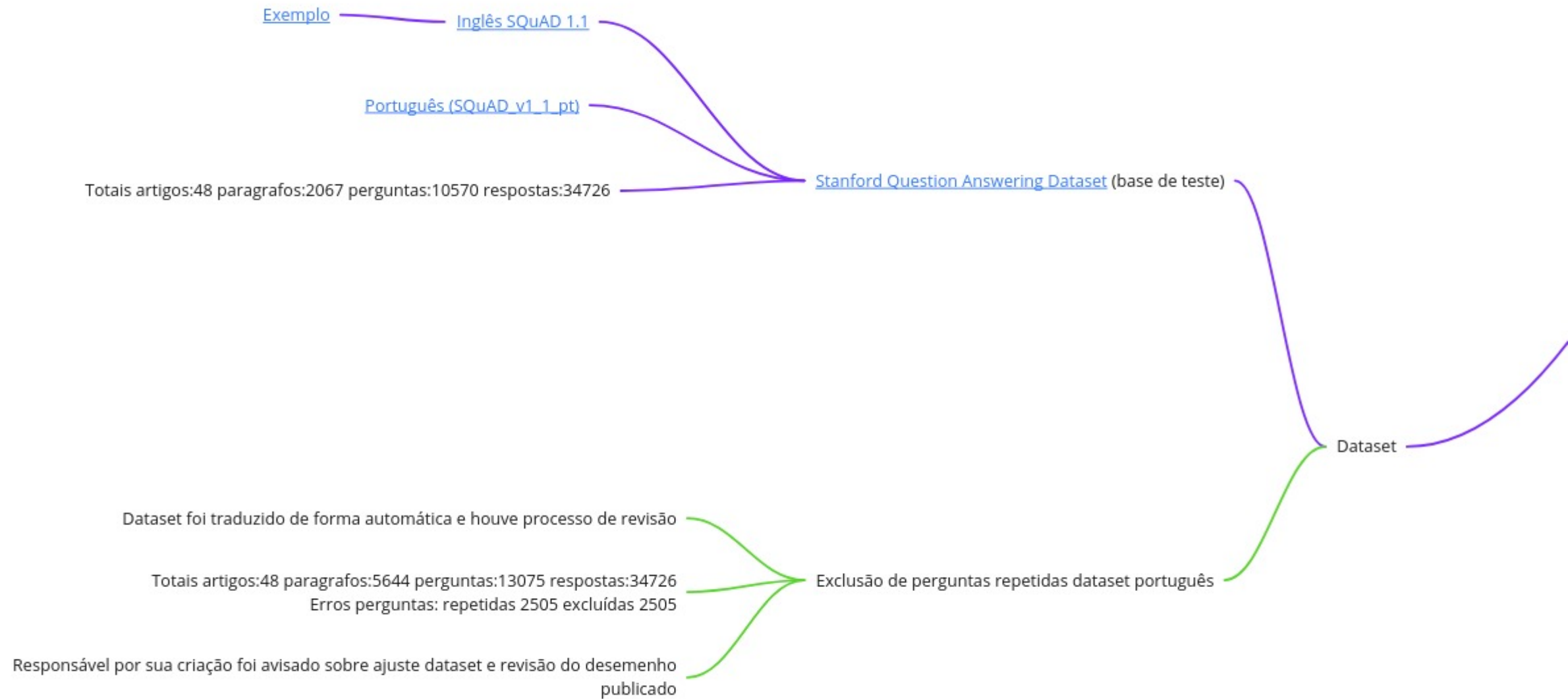
[GPT-neo-1.3B](https://huggingface.co/openai/gpt-neo-1.3B)

[GPT-Neo-2.7B](https://huggingface.co/openai/gpt-neo-2.7B)

[GPT-J_6B](https://huggingface.co/openai/gpt-j-6B)

<https://huggingface.co/EleutherAI>

Context Learning



Evaluations' rastro

Pandas Dataframe

Model

name_model

Learning Method

name

Squad Dataset

ind_language

Target

Context Information

num_question
datetime_execution
time_exec_total
time_exec_per_question

Parameters

Task

num_top_k = 3
num_max_answer_length= 80 tokens (context) / 160 letras (transfer)

Transfer

num_doc_stride = 128
if_handle_impossible_answer = False

Data

descr_filter = {}

Context

val_length_penalty = 0
val_temperature = 0.1
list_stop_words = [',', '\n', '!']
if_do_sample = False
cod_prompt_format *

Results

Total Metric

cod_metric
value

Metric per question

id_question
cod_metric
value

Prompt format

instrucao_pt = 'Instrução: Com base no texto abaixo, responda de forma sucinta à pergunta, evitando repetir palavras da pergunta:\n\n'

instrucao_en = 'Instruction: Based on the text below, answer the question succinctly, avoiding repeating words from the question:\n\n'

texto_questao_resposta_pt = 'Texto:{context}\n\nPergunta:{question}\nResposta:'

texto_questao_resposta_en = 'Text:{context}\n\nQuestion:{question}\nAnswer:'

exemplo1_pt = 'Exemplo:\n\nTexto: Marcus nasceu em 1980 e trabalha no TCU desde 2005.\n\nPergunta: Quem nasceu em 1980?\nResposta: Marcus\n\n'

exemplo1_en = 'Example:\n\nText: Marcus was born in 1980 and has worked at TCU since 2005.\n\nQuestion: Who was born in 1980?\nAnswer: Marcus\n\n' (...)

dict_prompt_format = {

1: {"prompt": instrucao_pt + texto_questao_resposta_pt, "num_shot":0, "ind_language":'pt'},

101: {"prompt": instrucao_en + texto_questao_resposta_en, "num_shot":0, "ind_language":'en'},

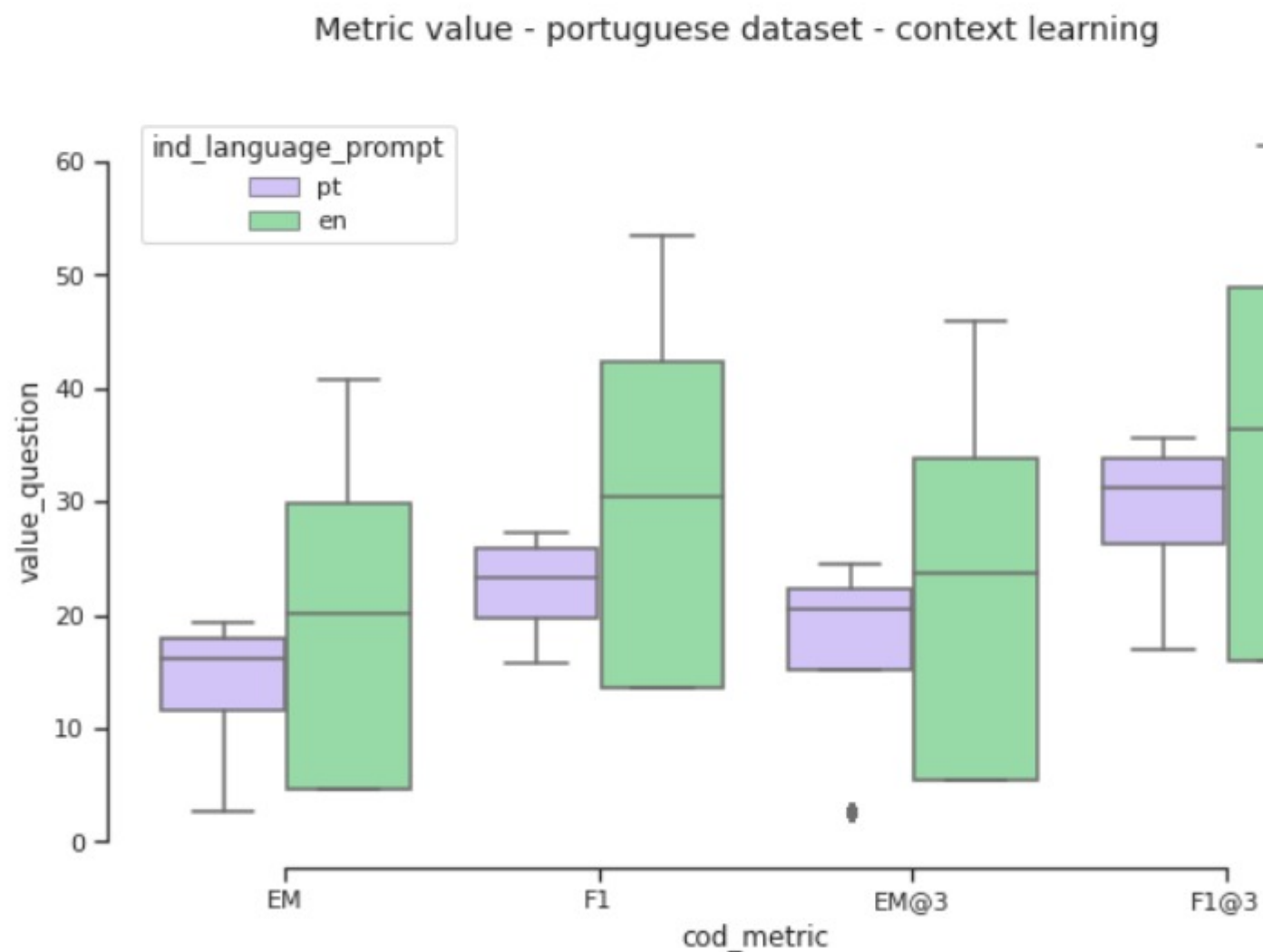
(...)

4: {"prompt": instrucao_pt + exemplo2_pt_tpp + texto_questao_resposta_pt, "num_shot":2, "format_example": "tpp", "ind_language":'pt'},

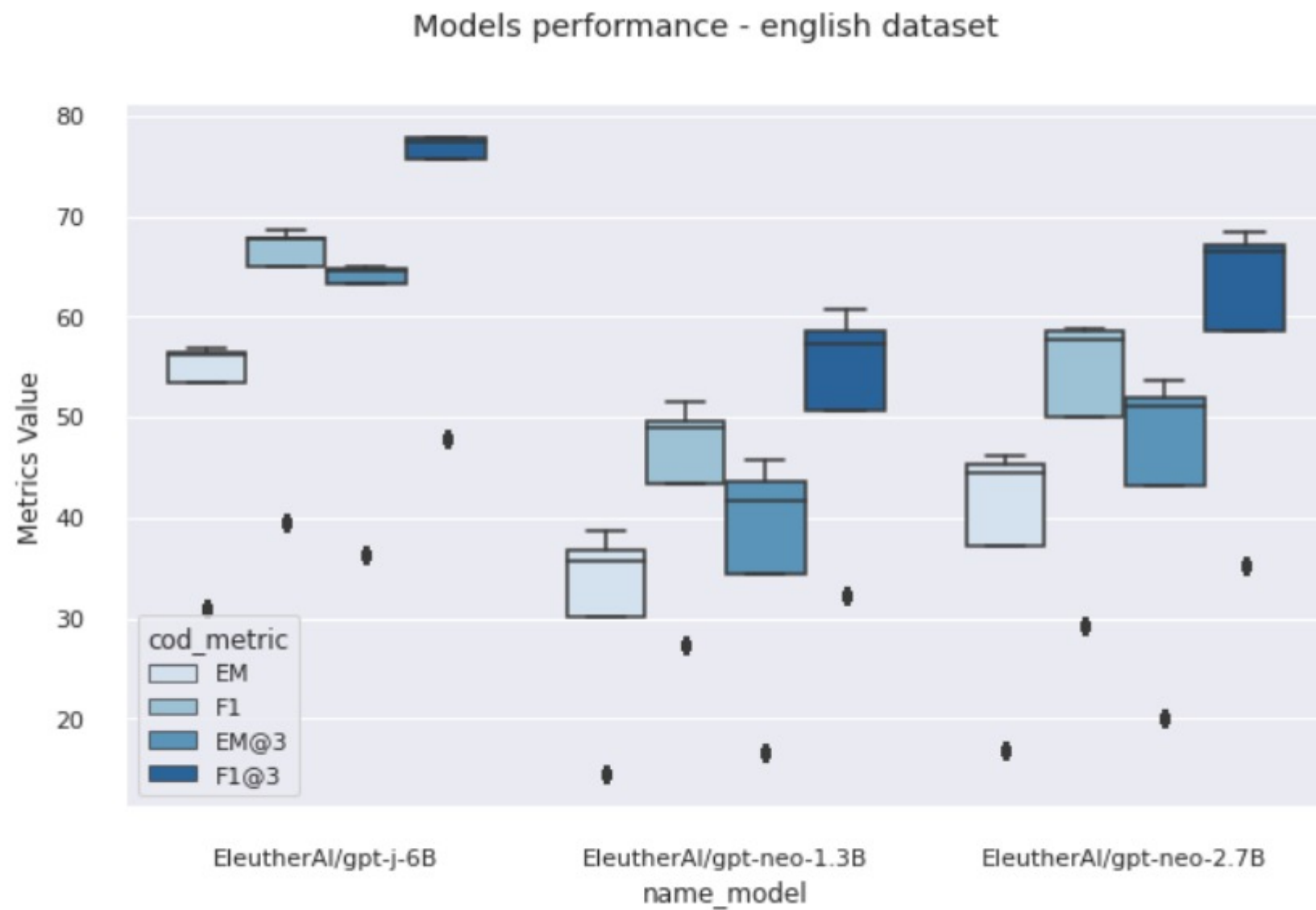
104: {"prompt": instrucao_en + exemplo2_en_tpp + texto_questao_resposta_en, "num_shot":2, "format_example": "tpp", "ind_language":'en'},

}

Alguns resultados alcançados - Context Learning



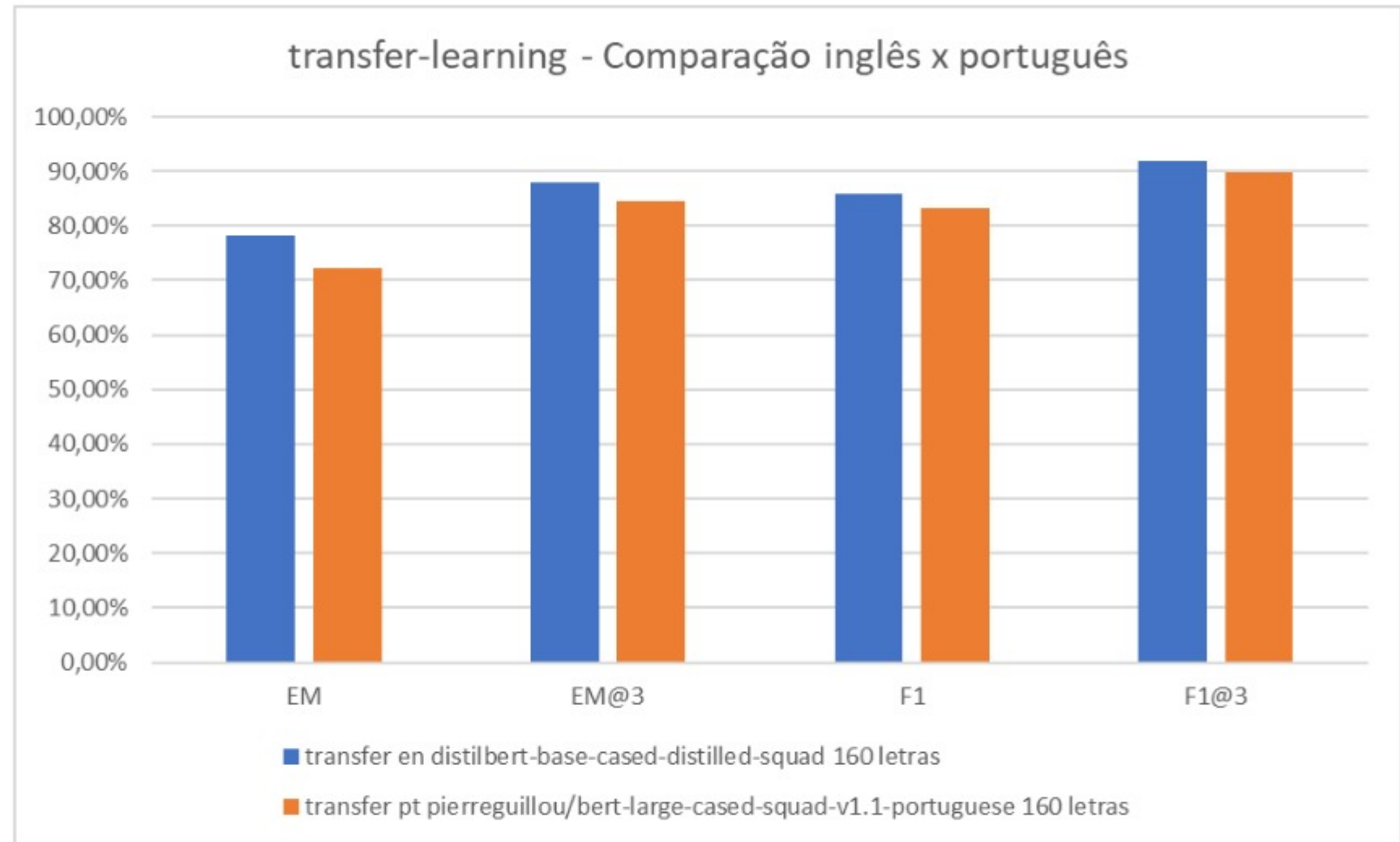
Context Learning



Context Learning



Transfer Learning



Tempo de execução

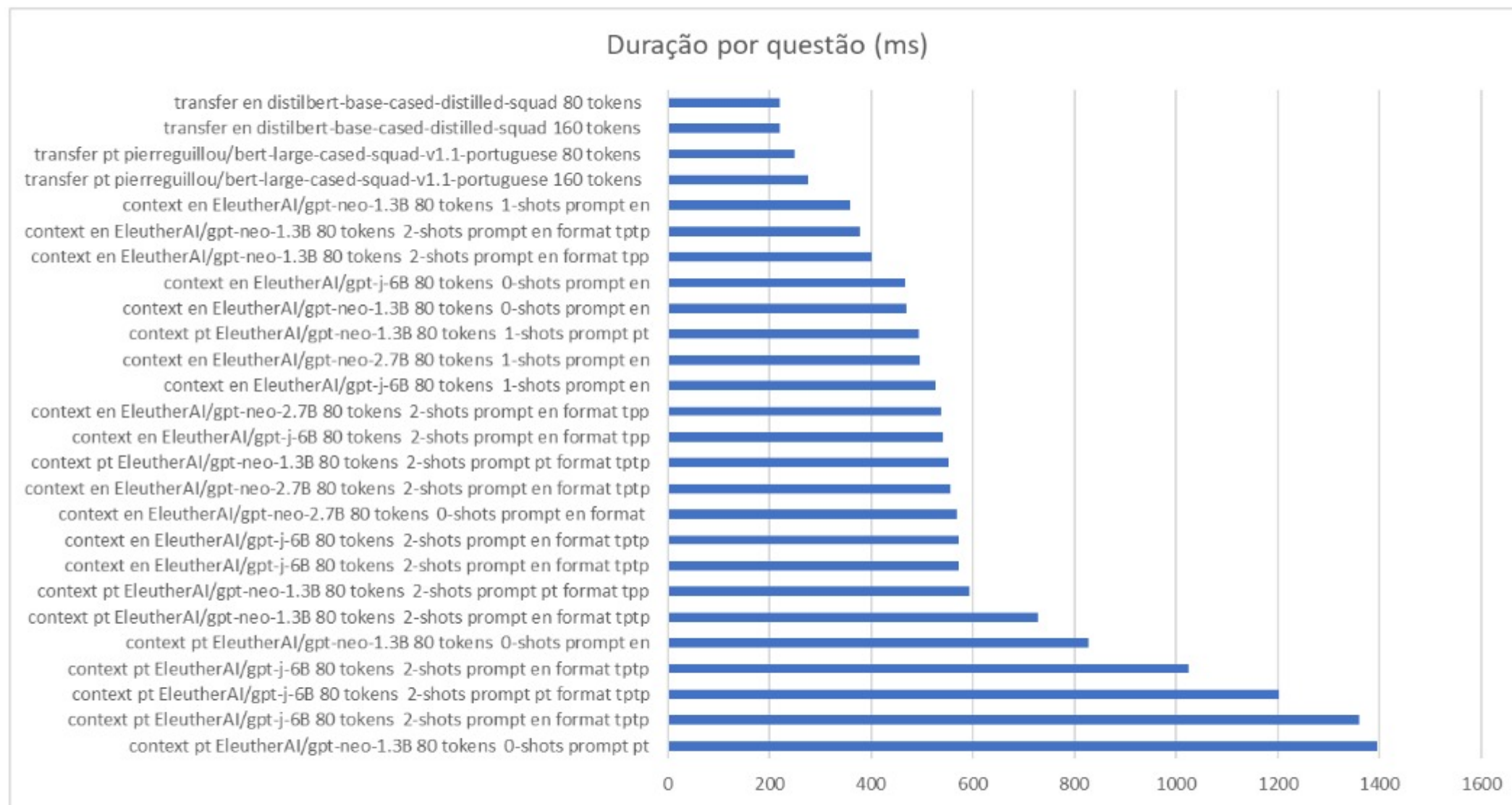
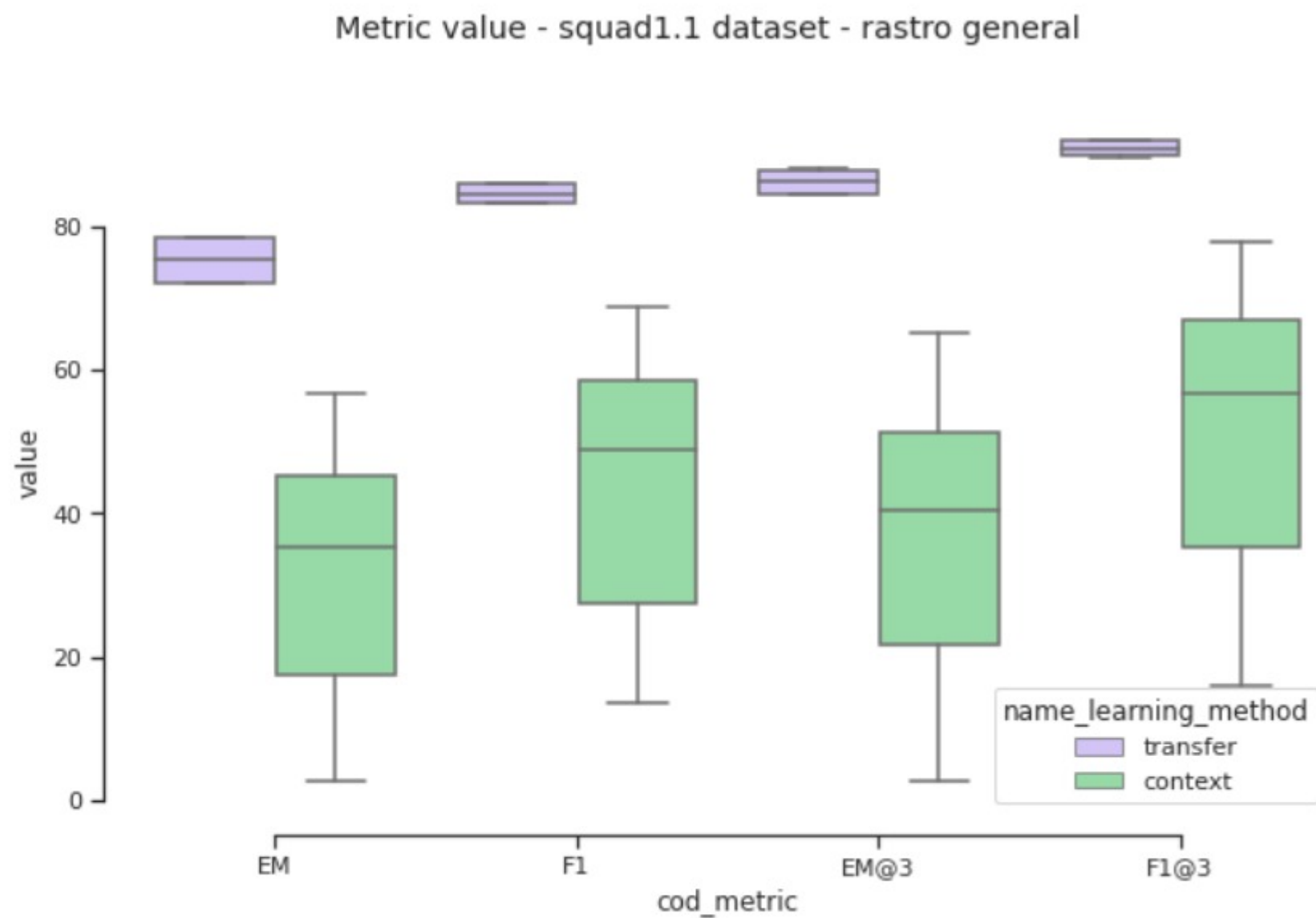


Figura 5 - Duração das execuções por questões, em milissegundos

Dá para comparar?



Trabalhos futuros

Repetir com GPT-3 (175B) e GPT-Neo-X (20B)

Avaliar dataset Squad 2.0 (possibilidade de não haver resposta)

Rever parâmetro usado indevidamente no context (pipe): min_length = 2

Para finalizar...

Comparing Performance of Learning Methods, which is better:
Context Learning and Transfer Learning?

Perguntando a quem entende...

Depende de condições de contexto
(inclusive de temperatura que
simula inspiração)
e de linguagem!

