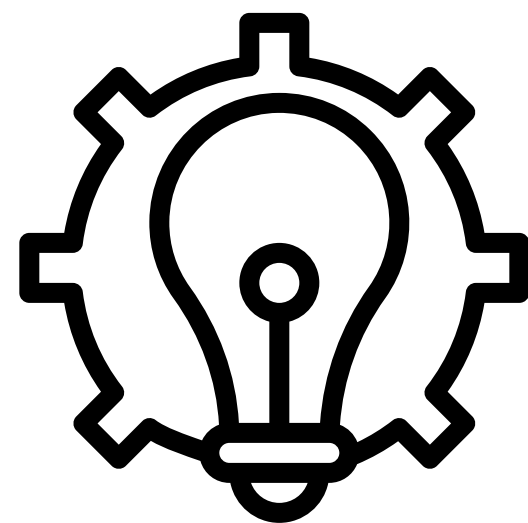


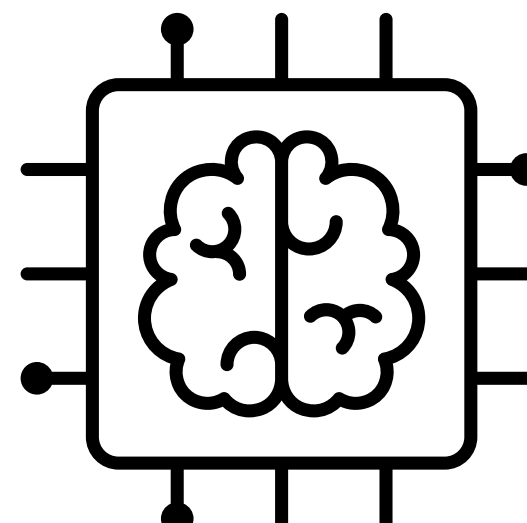
Modelo para o processamento de Linguagem Natural

Chat-BTG

Objetivo da Sprint



Implementação de
nova feature



Novas propostas de
Modelo PLN

AGENDA



Retrospectiva
Sprint 3



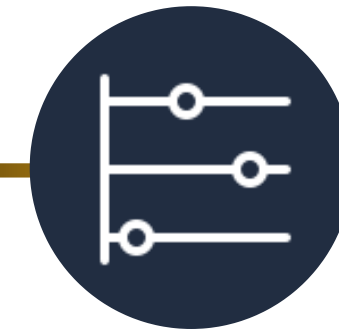
Arquitetura de
solução



Implementação de
nova feature



Novos modelos



Cronograma

01

Retrospectiva Sprint 3

Retrospectiva da Sprint 3



Lematização



Modelo Word2Vec



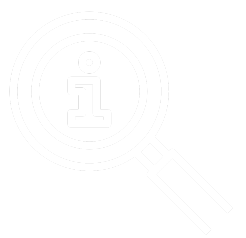
Resultados a partir das métricas

02

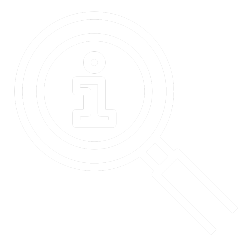
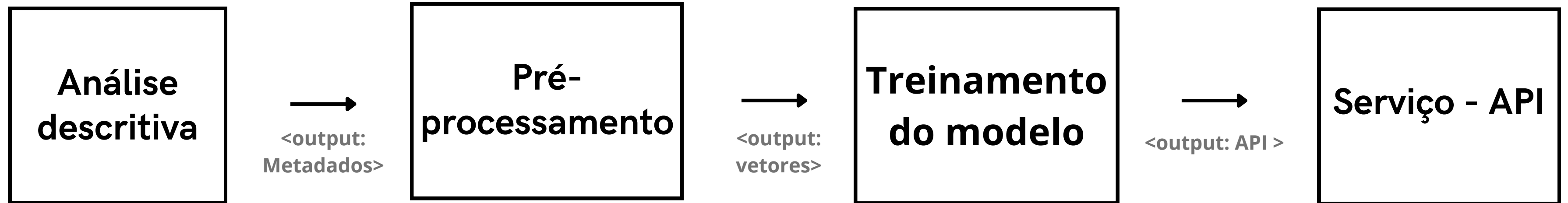
Arquitetura de solução

Extração, transporte e processamento

Externo



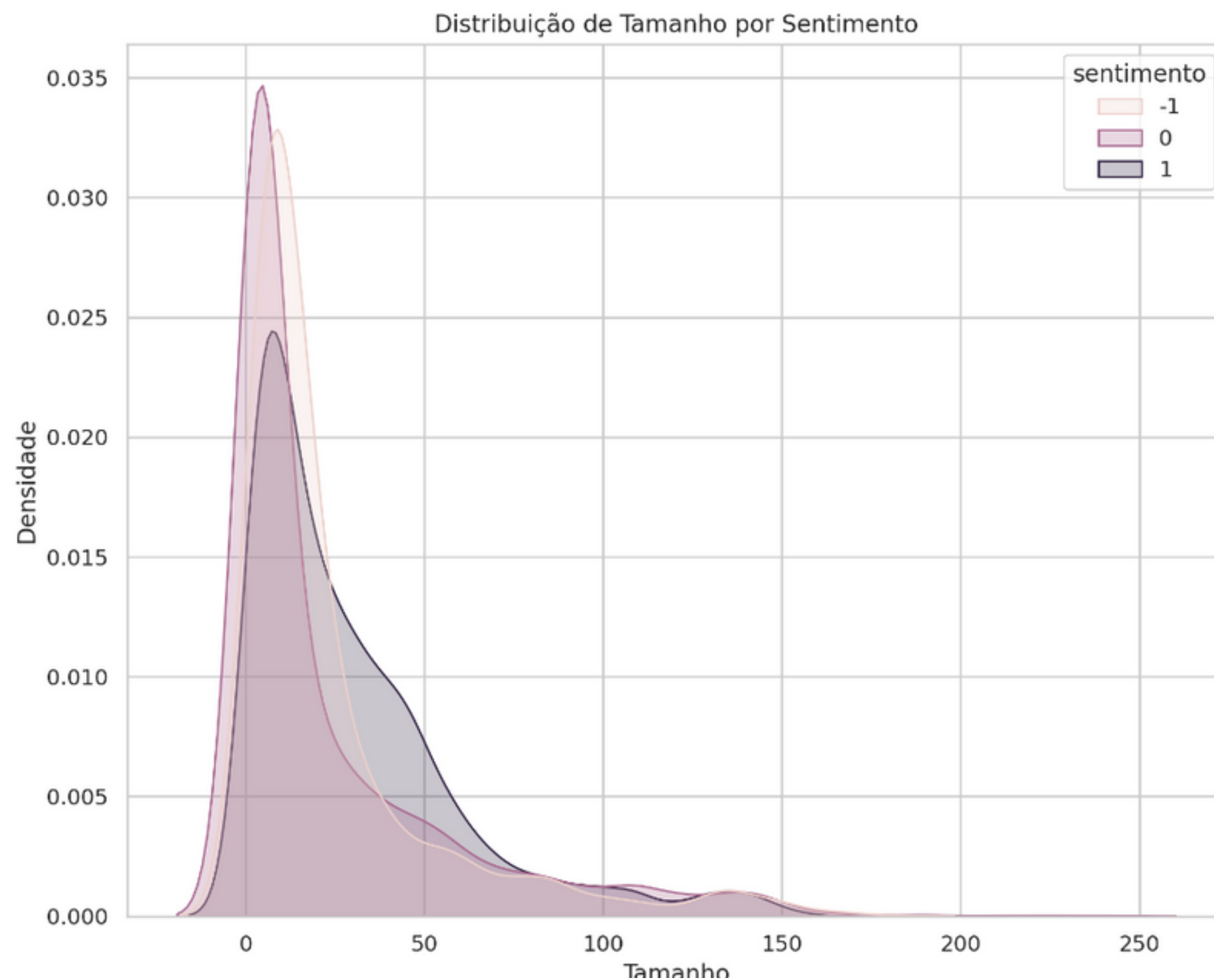
Arquitetura - Interno



03

Implementação de nova feature

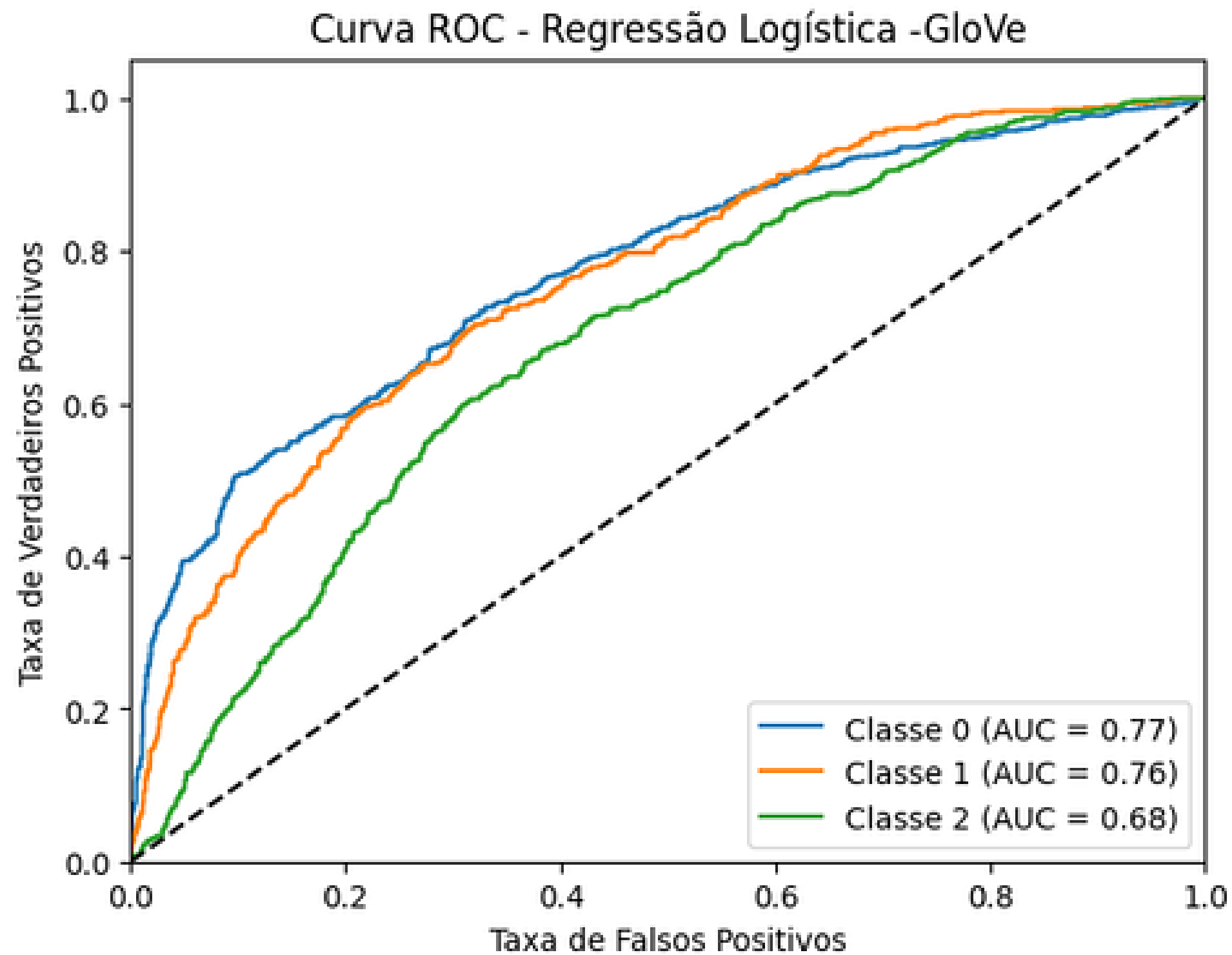
Quantidade de palavras por comentário



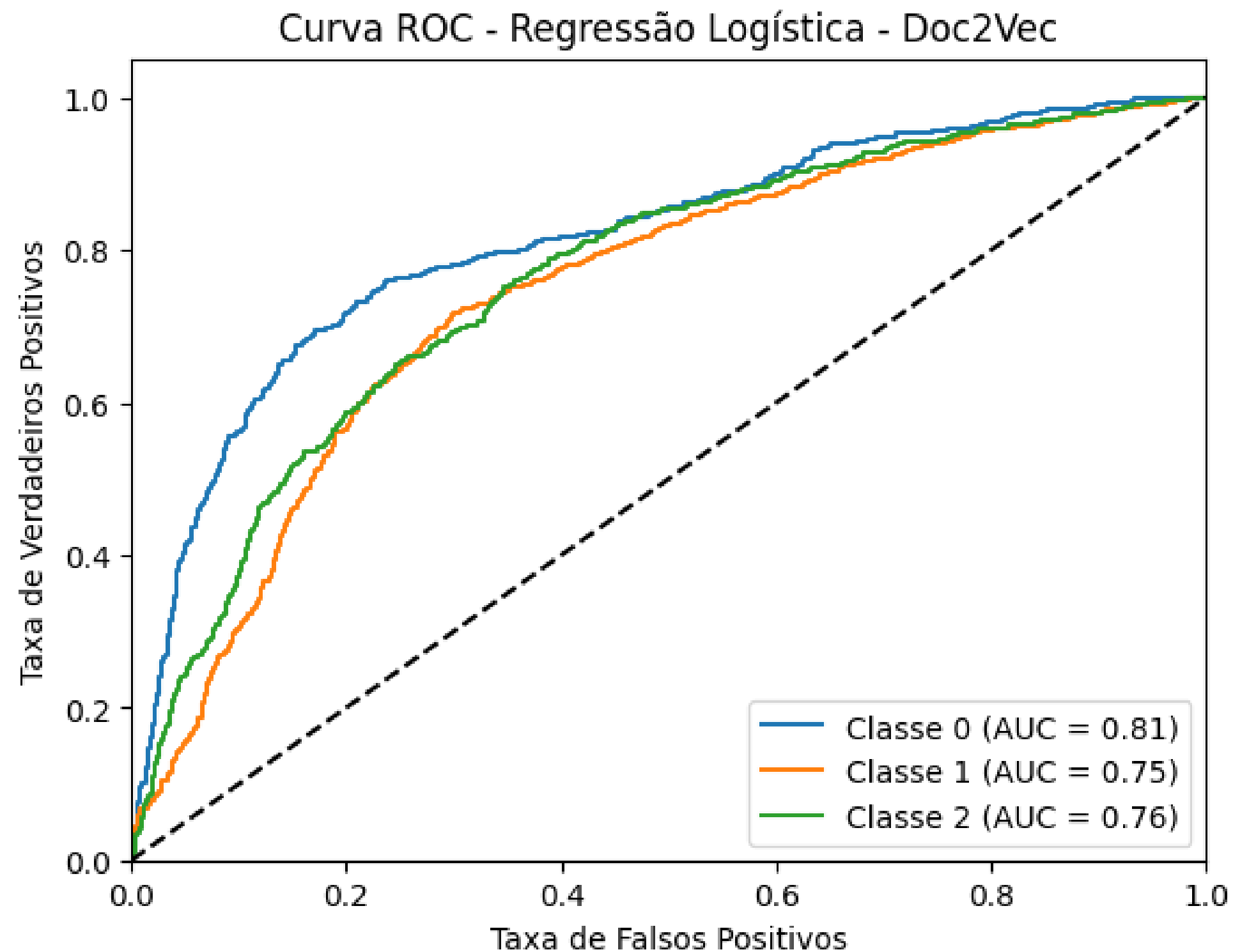
05

Novos modelos - Regressão Logística e Rede Neural

GloVe



Doc2Vec



BERT (Rede Neural)

	precision	recall	f1-score	<pre>{'input_ids': tensor([[101, 164, 112, ..., 0, 0, 0], [101, 164, 112, ..., 0, 0, 0], [101, 164, 112, ..., 0, 0, 0], ..., [101, 164, 112, ..., 0, 0, 0], [101, 164, 112, ..., 0, 0, 0], [101, 164, 112, ..., 0, 0, 0]]), 'token_type_ids': tensor([[0, 0, 0, ..., 0, 0, 0], [0, 0, 0, ..., 0, 0, 0], [0, 0, 0, ..., 0, 0, 0], ..., [0, 0, 0, ..., 0, 0, 0], [0, 0, 0, ..., 0, 0, 0], [0, 0, 0, ..., 0, 0, 0]]), 'attention_mask': tensor([[1, 1, 1, ..., 0, 0, 0], [1, 1, 1, ..., 0, 0, 0], [1, 1, 1, ..., 0, 0, 0], ..., [1, 1, 1, ..., 0, 0, 0], [1, 1, 1, ..., 0, 0, 0], [1, 1, 1, ..., 0, 0, 0]])}</pre>
NEGATIVE	0.63	0.78	0.70	
NEUTRAL	0.84	0.60	0.70	
POSITIVE	0.66	0.79	0.72	
accuracy			0.71	
macro avg	0.71	0.72	0.71	
weighted avg	0.73	0.71	0.71	

TF-IDF

Relatório de Classificação:

	precision	recall	f1-score	support
-1	0.76	0.57	0.65	479
0	0.73	0.83	0.78	1047
1	0.79	0.77	0.78	913

FastText



ELMo

```
[39] embeddings
```

```
array([[[ 0.6984197 ,  0.904021  ,  0.25575772, ..., -0.34985805,
          0.10431598,  1.0274081 ],
        [ 0.510448  ,  0.2128914 , -0.23980387, ..., -0.3364178 ,
          0.45042026,  0.7918285 ],
        [ 0.8487605 ,  0.6977814 , -0.44602975, ..., -0.41780114,
          0.34349352,  0.17138174],
        ...,
        [ 0.18144268,  1.1787145 , -0.43670553, ...,  0.4877972 ,
          0.02557464,  0.37615347],
        [ 0.23964994,  1.1514239 , -1.3816105 , ...,  0.01687086,
          0.3460997 , -0.8357353 ],
        [ 0.50669736,  0.59603035, -1.8193209 , ...,  0.15288985,
          1.1800405 , -0.7781176 ]]], dtype=float32)
```

Comparação entre modelos

Modelos	Recall	Ordem de assertividade
GloVe	0.74	2º
Doc2Vec	0.77	1º
BERT	0.73	3º
TF-IDF	0.72	4º
FastText	0.68	5º
ELMo	-	-

06

Cronograma

Etapa	Status	Sprint
Entendimento de Negócios	Feito	1
Mapeamento de Persona	Feito	1
Modelo de Bag of Words	Feito	2
Modelo utilizando Word2Vec e Análise de sentimento	Feito	3
Implementação de nova feature e desenvolvimento de novos modelos	Feito	4
Solução Final em formato MVP	A iniciar	5

Muito obrigado pela atenção!

Estamos disponíveis para maiores esclarecimentos.