In [1]:
```
# !pip install pycryptodome==3.15.0
# !pip install PyPDF2
```

In [2]:
```
# !python -m pip install spacy
# !python -m spacy download pt
```

In [3]:
```
import nltk
```

In [4]:
```
import numpy as np
```

In [5]:
```
from nltk.util import ngrams
```

In [6]:
```
import pandas as pd
```

In [7]:
```
from sklearn.model_selection import train_test_split
```

In [8]:
```
from sklearn.feature_extraction.text import CountVectorizer
```

In [9]:
```
from sklearn import linear_model
```

In [10]:
```
import string
```

In [11]:
```
from sklearn.linear_model import LogisticRegression
```

In [12]:
```
from unidecode import unidecode
```

In [13]:
```
from sklearn.feature_extraction.text import TfidfVectorizer
```

In [ ]:
```
import matplotlib.pyplot as plt
```

In [14]:
```
perguntas = [
    'Como esse COIL afetou sua consciência global?',
    'Que novas perspectivas você descobriu durante sua experiência no COIL?',
    'Que pontos em comum você notou durante sua experiência com o COIL?',
    'Você recomendaria uma experiência COIL a um amigo? Por que sim ou por que não?
    'O que você diria que os alunos precisam fazer se quiserem ter uma experiência
    'Se você fosse fazer o COIL com colegas internacionais novamente, o que faria d
    'O que tornou a experiência COIL única ou especial?',
    'O que você aprendeu sobre seus colegas internacionais? O que você aprendeu sob
    'De modo geral, qual foi o momento de aprendizado mais interessante e útil para
    'O que você aprendeu com o COIL que pode ajudá-lo em outras áreas de sua vida?'
    'O que você aprendeu com a apresentação dos seus colegas de equipe do Brasil nc
    'Seus aprendizados do'
]
```

respostas = ['''- O COIL abriu meus olhos para o quanto os alunos brasileiros são ótimos. Desde o início, fiquei agradavelmente surpreso com a qualidade do inglês de todos, porque eu estava prevendo que haveria um desafio na comunicação. Supondo que os brasileiros geralmente usam o português como idioma principal. Eles me provaram que seu nível de inglês é tão bom quanto o de qualquer falante nativo de inglês.''', '''- Depois de conhecer todos por meio do Padlet, adoro o fato de que muitos de nós viemos de diferentes partes do mundo, passando por diferentes fases da vida, com nossos próprios objetivos e sonhos, e ainda assim nos cruzamos. Embora eu não conheça todos pessoalmente, é bom conhecer as pessoas nas fotos por meio de suas palavras. Até mesmo os detalhes mais simples, como seus programas/filmes favoritos, despertaram meu interesse em pesquisá-los.''', '''- Os pontos em comum que notei são que todos nós gostamos de sair e compartilhamos hobbies semelhantes. Como assistir a programas da Netflix, dançar nossas músicas favoritas, ler livros, etc. .... Percebi que uma aluna da FATEC, Letícia, mencionou que seu artista favorito é Bruno Mars e

que ele também é um dos principais artistas da minha lista de reprodução do Spotify. Outra aluna, Kamila, mencionou que gosta de assistir a dramas coreanos como um de seus passatempos. Eu assisto a dramas coreanos há muito tempo e é bom ver o reconhecimento internacional que eles estão recebendo. Em suma, se Kamilla e eu fôssemos colegas de classe, eu com certeza pediria a ela que compartilhasse seus dramas coreanos favoritos.''', '''- Eu recomendaria a um amigo, pois é uma atividade empolgante em comparação com os trabalhos normais em grupo. Vejo isso como uma oportunidade de se expor a colegas de outro país e conhecer sua cultura. Não pode ser comparado ao trabalho com nossos próprios alunos estrangeiros, pois eles estão adaptados ao funcionamento do nosso sistema escolar. A experiência do COIL também nos permite aprender com perspectivas diferentes e adquirir conhecimentos que não são comuns no Canadá.''', '''- Comunicação e manter a responsabilidade mútua. Pessoalmente, não sou fã de trabalho em grupo, pois a carga de trabalho nunca é distribuída de forma homogênea. No entanto, meus colegas de equipe nesse projeto foram ótimos. Todos contribuíram para o projeto, o que significa que nenhuma pessoa ficou carregando um peso maior sobre os ombros. Meus colegas de equipe também apoiaram muito uns aos outros e proporcionaram um espaço confortável para perguntas e ajuda adicional.''', '''- Gostaria de trabalhar com eles em uma tarefa em vez de compartilhar nossas próprias respostas no SLACK. O fato de os dois países trabalharem separadamente no mesmo exercício anulou o objetivo da colaboração. Além disso, eu gostaria de trabalhar em um tipo diferente de projeto que incentivasse mais a comunicação, como uma apresentação em PowerPoint.''', '''- O que tornou a experiência do COIL especial foi a oportunidade de colaborar com outra escola. Com desafios como diferença de fuso horário e barreiras linguísticas, não tivemos problemas, principalmente por causa da incrível capacidade de nossos colegas internacionais. Além disso, a tecnologia moderna ajuda a todos com uma comunicação tranquila que nos permite conectar por meio do SLACK e do WhatsApp.''', '''- Foi incrível trabalhar com nossa colega internacional, Valerie. Por ser a única pessoa internacional, ela trabalhava com mais frequência em nosso horário. Seu inglês também era excelente, portanto não havia barreira de idioma entre nós quando se tratava de comunicação. Embora não fôssemos do mesmo país, eu não conseguia perceber a diferença, pois ela era muito boa em suas respostas. - Algo que aprendi sobre mim mesmo é que estou ansioso para conhecer e trabalhar com novas pessoas. Quando nosso professor nos apresentou o projeto, eu estava bastante ansioso por causa de coisas como diferença de fuso horário, barreiras linguísticas, etc. No entanto, uma parte de mim também estava muito animada para conhecer alunos de outros países. Para saber mais sobre o que eles aprendem na escola ou como é o programa de RH deles. Aprendi que as políticas de RH deles não são muito diferentes das nossas.''', '''- O momento mais interessante para mim ao trabalhar com Valerie é que verificávamos o processo uma da outra, apesar de não estarmos trabalhando no mesmo trabalho. Isso mostrou que ainda estávamos fazendo pequenas coisas como uma equipe. Eu não esperava por isso, mas ela foi a primeira a tomar a iniciativa. Não posso elogiar Valerie o suficiente, pois ela deixou uma ótima impressão. Eu teria adorado conhecê-la melhor, mas nós duas estávamos ocupadas com nossas agendas.''', '''- Minha maior lição disso tudo é tentar ser mais aberto e aceitar novos desafios. Como sou introvertido, fico muito ansioso ao conhecer novas pessoas e normalmente mantenho os colegas de escola à distância. Estar no RH significa que trabalharei com pessoas em um nível pessoal e isso é algo que ainda estou aprendendo a ajustar.''', '''- Fiquei sabendo que minha colega de equipe, Valerie, é uma mulher casada que está matriculada em um programa de RH no Brasil. Ela tem um marido amoroso e uma cachorra chamada Tiffany. Ela é fã de Whitney Houston e até tocou "I Will Always Love You" em seu casamento. Ela também mencionou que adora frutas e legumes, o que me sugere que ela tem um estilo de vida bastante saudável.''', '''- O que aprendi no Módulo 4 de Valerie é que ela trabalha em casa, o que torna sua experiência um pouco diferente da minha. No que diz respeito à comunicação, é mais difícil, pois você não tem tempo para se relacionar com seus colegas de trabalho individualmente. Em termos de fatores de estresse, compartilhamos muitos fatores semelhantes, como longas horas de trabalho e liderança deficiente. Em geral, compartilhamos experiências semelhantes, apesar de virmos de diferentes partes do mundo, quando se trata de locais de trabalho. Os poucos motivos que diferenciam nossas experiências são coisas como nossa cultura ou leis de trabalho diferentes.''']

In [16]:
```
# print(f"Tamanho Perguntas: {len(perguntas)}   Respostas: {len(respostas)}")
```

```
Tamanho Perguntas: 12   Respostas: 12
```

In [1]:
```
# Carregar a base de dados de perguntas e respostas
df_respostas = pd.read_csv("../../../dados/pesquisa-sol/reflexoes-coil.csv", encodi
    names=['id', 'hora_inicio', 'hora_conclusao', 'email', 'nome', 'ult_mod', 'nome
           'q5', 'q6', 'q7', 'q8', 'q9', 'q10'], header=1)
df_respostas.describe()
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[1], line 2
      1 # Carregar a base de dados de perguntas e respostas
----> 2 df = pd.read_csv("../../../dados/pesquisa-sol/reflexoes-coil.csv", encodin
g='utf8', delimiter=";", quotechar='"',
      3     names=['id', 'hora_inicio', 'hora_conclusao', 'email', 'nome', 'ult_mo
d', 'nome_entrevistado', 'q1', 'q2', 'q3', 'q4',
      4            'q5', 'q6', 'q7', 'q8', 'q9', 'q10'], header=1)
      5 df.describe()

NameError: name 'pd' is not defined
```

In [17]:
```
# Carregar a base de dados para treino
df = pd.read_csv("../../../dados/imdb-reviews-pt-br.csv")
df.describe()
```

Out[17]:

	id
count	49459.000000
mean	24730.960917
std	14277.792868
min	1.000000
25%	12366.500000
50%	24731.000000
75%	37095.500000
max	49460.000000

In [ ]:
```
# Remover colunas inúteis da base de respostas
df_respostas.drop(
    columns = ['hora_inicio', 'hora_conclusao', 'email', 'nome', 'ult_mod'],
    inplace = True
)
```

Preparo do data set

In [18]:
```
# Remover colunas inúteis
df.drop(
    columns = ["text_en"],
    inplace = True
)
```

In [19]:
```
# Remover linhas com conteudo N/A
df.dropna(inplace = True)
```

Criar a coluna numerica para representar o 'pos' e 'neg'

In [20]:
```
df = df.assign(sentiment_value = [1 if i == 'pos' else 0  for i in df["sentiment"]]
```

In [21]:
```
df
```

Out[21]:

	id	text_pt	sentiment	sentiment_value
0	1	Mais uma vez, o Sr. Costner arrumou um filme p...	neg	0
1	2	Este é um exemplo do motivo pelo qual a maiori...	neg	0
2	3	Primeiro de tudo eu odeio esses raps imbecis, ...	neg	0
3	4	Nem mesmo os Beatles puderam escrever músicas ...	neg	0
4	5	Filmes de fotos de latão não é uma palavra apr...	neg	0
...	...	...	...	...
49454	49456	Como a média de votos era muito baixa, e o fat...	pos	1
49455	49457	O enredo teve algumas reviravoltas infelizes e...	pos	1
49456	49458	Estou espantado com a forma como este filme e ...	pos	1
49457	49459	A Christmas Together realmente veio antes do m...	pos	1
49458	49460	O drama romântico da classe trabalhadora do di...	pos	1

49459 rows × 4 columns

Preparo do texto text_pt

Funções para o tratamento de texto

In [22]:
```
# Função para preparar as palavras do texto
def preparar_texto( texto ):
    # Converter em minusculo
    texto_limpo = texto.lower()
    # Remover pontuação e caracteres especiais
    translator = str.maketrans('\n\r\t', '   ', string.punctuation)
    texto_limpo = texto_limpo.translate(translator)
    # Remover os acentos
    texto_limpo = unidecode(texto_limpo)
    return texto_limpo
```

In [23]:
```
# Remover Stopwords
stopwords = nltk.corpus.stopwords.words('portuguese')
def remove_stopwords( texto ):
    tokens = []
    for token in nltk.tokenize.word_tokenize(texto, language='portuguese'):
        if token not in stopwords:
            tokens.append(token)
    return " ".join(tokens)
```

In [42]:
```
len(stopwords)
```

Out[42]:
```
207
```

In [25]:
```
# Lematização
def lematizar( texto ):
    doc = nlp(texto)
    doc_lemma = [token.lemma_ for token in doc if token.pos_ == 'NOUN']
    return " ".join(doc_lemma)
```

In [26]:
```
# Stemming
stemmer = nltk.RSLPStemmer()
```

```
def stemmer_text( texto ):
    if isinstance(texto, str):
        lista_palavras = texto.split(" ")
        nova_lista = []
        for palavra in lista_palavras:
            stemmed = stemmer.stem( palavra )
            nova_lista.append(stemmed)
        return " ".join(nova_lista)
    else:
        return texto
```

In [27]:
```
def transformar_texto( texto ):
    texto_limpo = preparar_texto( texto )
    return texto_limpo
    # texto_sem_stopwords = remove_stopwords(texto_limpo)
    # texto_stemmed = stemmer_text(texto_sem_stopwords)
    # return texto_stemmed
```

Prepara o texto e coloca em uma variavel separada

In [28]:
```
count = 0
def invocar_transformar_texto( texto ):
    global count
    count += 1
    if count % 1000 == 0:
        print(f"Analisando linha: {count}")
    return transformar_texto( texto )

texto_preparado = df["text_pt"].apply(invocar_transformar_texto)
```

```
Analisando linha: 1000
Analisando linha: 2000
Analisando linha: 3000
Analisando linha: 4000
Analisando linha: 5000
Analisando linha: 6000
Analisando linha: 7000
Analisando linha: 8000
Analisando linha: 9000
Analisando linha: 10000
Analisando linha: 11000
Analisando linha: 12000
Analisando linha: 13000
Analisando linha: 14000
Analisando linha: 15000
Analisando linha: 16000
Analisando linha: 17000
Analisando linha: 18000
Analisando linha: 19000
Analisando linha: 20000
Analisando linha: 21000
Analisando linha: 22000
Analisando linha: 23000
Analisando linha: 24000
Analisando linha: 25000
Analisando linha: 26000
Analisando linha: 27000
Analisando linha: 28000
Analisando linha: 29000
Analisando linha: 30000
Analisando linha: 31000
Analisando linha: 32000
Analisando linha: 33000
Analisando linha: 34000
Analisando linha: 35000
Analisando linha: 36000
Analisando linha: 37000
Analisando linha: 38000
Analisando linha: 39000
Analisando linha: 40000
Analisando linha: 41000
Analisando linha: 42000
Analisando linha: 43000
Analisando linha: 44000
Analisando linha: 45000
Analisando linha: 46000
Analisando linha: 47000
Analisando linha: 48000
Analisando linha: 49000
```

Criar o dicionario e o Bag of Words

In [29]:
```
# # Criar o bag_of_words com base no CountVectorizer
# vetorizador = CountVectorizer(max_features=100, lowercase=False)
# bag_of_words = vetorizador.fit_transform(texto_preparado)
# bag_of_words.shape    # Verifica o formato da Matriz
```

In [30]:
```
# # Criar o bag_of_words com base no TfidfVectorizer usando NGram 3
# vetorizador = TfidfVectorizer(lowercase=False, max_features=100, ngram_range=(1,
# bag_of_words = vetorizador.fit_transform(texto_preparado)
# bag_of_words.shape
```

In [31]:
```
# Criar o bag_of_words com base no TfidfVectorizer
vetorizador = TfidfVectorizer(lowercase=False, max_features=2000)
bag_of_words = vetorizador.fit_transform(texto_preparado)
bag_of_words.shape
```

Out[31]:
```
(49459, 2000)
```

In [32]:
```
dicionario = vetorizador.get_feature_names_out()
```

In [33]:
```
bow = pd.DataFrame.sparse.from_spmatrix(bag_of_words, columns=dicionario)
```

In [34]:
```
train, test, train_class, test_class = train_test_split(bow,
                                                        df["sentiment_value"],
                                                        random_state = 100)
```

In [35]:
```
reg_logistica = LogisticRegression()
reg_logistica.fit(train, train_class)
acuracia = reg_logistica.score(test, test_class)
```

In [36]:
```
acuracia
```

Out[36]:
```
0.874646178730287
```

Acuracias

Sem tratamento do texto ==> 0.7013344116457744

Colocando as palavras em minusculo ==> 0.7107966033158107

Remover os caracteres especiais ==> 0.7273756570966438

Remover os acentos ==> 0.7272139102304893

Remover as stopwords ==> 0.7199353012535382

Remover stopwords e aplicar raiz do texto (stemm) ==> 0.7532551556813587

Aplicando o TF_ID ao inves da contagem de palavras ==> 0.7553578649413668

Aplicando NGRAM (1,3) ==> 0.7553578649413668

Aplicando TF_ID e MAX_Features = 200 ==> 0.7955519611807521

Aplicando TF_ID e MAX_Features = 300 ==> 0.8188435099069955

Aplicando TF_ID e MAX_Features = 500 ==> 0.8418924383340073

Aplicando TF_ID e MAX_Features = 1000 ==> 0.8639708855640922

Aplicando TF_ID e MAX_Features = 2000 ==> 0.8760210270926001

In [37]:
```
# Tratar as respostas
respostas_preparadas = np.array([invocar_transformar_texto(r) for r in respostas])
```

In [38]:
```
respostas_preparadas
```

Out[38]:
```
array([' o coil abriu meus olhos para o quanto os alunos brasileiros sao otimos de
sde o inicio fiquei agradavelmente surpreso com a qualidade do ingles de todos por
que eu estava prevendo que haveria um desafio na comunicacao supondo que os brasil
eiros geralmente usam o portugues como idioma principal eles me provaram que seu n
ivel de ingles e tao bom quanto o de qualquer falante nativo de ingles',
       ' depois de conhecer todos por meio do padlet adoro o fato de que muitos de
nos viemos de diferentes partes do mundo passando por diferentes fases da vida com
nossos proprios objetivos e sonhos e ainda assim nos cruzamos embora eu nao conhec
a todos pessoalmente e bom conhecer as pessoas nas fotos por meio de suas palavras
ate mesmo os detalhes mais simples como seus programasfilmes favoritos despertaram
meu interesse em pesquisalos',
       ' os pontos em comum que notei sao que todos nos gostamos de sair e compart
ilhamos hobbies semelhantes como assistir a programas da netflix dancar nossas mus
icas favoritas ler livros etc  percebi que uma aluna da fatec leticia mencionou qu
e seu artista favorito e bruno mars e que ele tambem e um dos principais artistas
da minha lista de reproducao do spotify outra aluna kamila mencionou que gosta de
assistir a dramas coreanos como um de seus passatempos eu assisto a dramas coreano
s ha muito tempo e e bom ver o reconhecimento internacional que eles estao receben
do em suma se kamilla e eu fossemos colegas de classe eu com certeza pediria a ela
que compartilhasse seus dramas coreanos favoritos',
       ' eu recomendaria a um amigo pois e uma atividade empolgante em comparacao
com os trabalhos normais em grupo vejo isso como uma oportunidade de se expor a co
legas de outro pais e conhecer sua cultura nao pode ser comparado ao trabalho com
nossos proprios alunos estrangeiros pois eles estao adaptados ao funcionamento do
nosso sistema escolar a experiencia do coil tambem nos permite aprender com perspe
ctivas diferentes e adquirir conhecimentos que nao sao comuns no canada',
       ' comunicacao e manter a responsabilidade mutua pessoalmente nao sou fa de
trabalho em grupo pois a carga de trabalho nunca e distribuida de forma homogenea
no entanto meus colegas de equipe nesse projeto foram otimos todos contribuiram pa
ra o projeto o que significa que nenhuma pessoa ficou carregando um peso maior sob
re os ombros meus colegas de equipe tambem apoiaram muito uns aos outros e proporc
ionaram um espaco confortavel para perguntas e ajuda adicional',
       ' gostaria de trabalhar com eles em uma tarefa em vez de compartilhar nossa
s proprias respostas no slack o fato de os dois paises trabalharem separadamente n
o mesmo exercicio anulou o objetivo da colaboracao alem disso eu gostaria de traba
lhar em um tipo diferente de projeto que incentivasse mais a comunicacao como uma
apresentacao em powerpoint',
       ' o que tornou a experiencia do coil especial foi a oportunidade de colabor
ar com outra escola com desafios como diferenca de fuso horario e barreiras lingui
sticas nao tivemos problemas principalmente por causa da incrivel capacidade de no
ssos colegas internacionais alem disso a tecnologia moderna ajuda a todos com uma
comunicacao tranquila que nos permite conectar por meio do slack e do whatsapp',
       ' foi incrivel trabalhar com nossa colega internacional valerie por ser a u
nica pessoa internacional ela trabalhava com mais frequencia em nosso horario seu
ingles tambem era excelente portanto nao havia barreira de idioma entre nos quando
se tratava de comunicacao embora nao fossemos do mesmo pais eu nao conseguia perce
ber a diferenca pois ela era muito boa em suas respostas  algo que aprendi sobre m
im mesmo e que estou ansioso para conhecer e trabalhar com novas pessoas quando no
sso professor nos apresentou o projeto eu estava bastante ansioso por causa de coi
sas como diferenca de fuso horario barreiras linguisticas etc no entanto uma parte
de mim tambem estava muito animada para conhecer alunos de outros paises para sabe
r mais sobre o que eles aprendem na escola ou como e o programa de rh deles aprend
i que as politicas de rh deles nao sao muito diferentes das nossas',
       ' o momento mais interessante para mim ao trabalhar com valerie e que verif
icavamos o processo uma da outra apesar de nao estarmos trabalhando no mesmo traba
lho isso mostrou que ainda estavamos fazendo pequenas coisas como uma equipe eu na
o esperava por isso mas ela foi a primeira a tomar a iniciativa nao posso elogiar
valerie o suficiente pois ela deixou uma otima impressao eu teria adorado conhecel
a melhor mas nos duas estavamos ocupadas com nossas agendas',
       ' minha maior licao disso tudo e tentar ser mais aberto e aceitar novos des
afios como sou introvertido fico muito ansioso ao conhecer novas pessoas e normalm
ente mantenho os colegas de escola a distancia estar no rh significa que trabalhar
ei com pessoas em um nivel pessoal e isso e algo que ainda estou aprendendo a ajus
tar',
```

```
       ' fiquei sabendo que minha colega de equipe valerie e uma mulher casada que
esta matriculada em um programa de rh no brasil ela tem um marido amoroso e uma ca
chorra chamada tiffany ela e fa de whitney houston e ate tocou i will always love
you em seu casamento ela tambem mencionou que adora frutas e legumes o que me suge
re que ela tem um estilo de vida bastante saudavel',
       ' o que aprendi no modulo 4 de valerie e que ela trabalha em casa o que tor
na sua experiencia um pouco diferente da minha no que diz respeito a comunicacao e
mais dificil pois voce nao tem tempo para se relacionar com seus colegas de trabal
ho individualmente em termos de fatores de estresse compartilhamos muitos fatores
semelhantes como longas horas de trabalho e lideranca deficiente em geral comparti
lhamos experiencias semelhantes apesar de virmos de diferentes partes do mundo qua
ndo se trata de locais de trabalho os poucos motivos que diferenciam nossas experi
encias sao coisas como nossa cultura ou leis de trabalho diferentes'],
      dtype='<U880')
```

In [39]:
```
vetorizador = TfidfVectorizer(lowercase=False, max_features=2000, vocabulary=dicior
bag_of_words_respostas = vetorizador.fit_transform(respostas_preparadas)
bag_of_words_respostas.shape
```

Out[39]:
```
(12, 2000)
```

In [40]:
```
reg_logistica.predict(bag_of_words_respostas)
```

```
D:\usr\anaconda3\lib\site-packages\sklearn\base.py:420: UserWarning: X does not ha
ve valid feature names, but LogisticRegression was fitted with feature names
  warnings.warn(
```

Out[40]:
```
array([1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1], dtype=int64)
```

In [41]:
```
# i = 0
# print("Linhas size: ", len(linhas))
# while i < len(linhas):
#     linha = linhas[i].lower()
#     linha_tokens = nltk.tokenize.word_tokenize(linha, language='portuguese')
#     linha_tokens_limpos = remove_stop_words(linha_tokens)
#     i += 1
#     if linha_tokens_limpos is None or len(linha_tokens_limpos) < 3:
#         continue
#     linha_ngrams = ngrams(linha_tokens_limpos, 3)
#     linha_points = 0
#     for pergunta in perguntas:
#         pergunta_tokens = nltk.tokenize.word_tokenize(pergunta, language='portugu
#         pergunta_no_stopwords = remove_stop_words(pergunta_tokens)
#         for ngram_tuple in linha_ngrams:
#             ngram_tupla_texto = " ".join(ngram_tuple)
#             if ngram_tupla_texto in pergunta.lower():
#                 linha_points += 1
# #                 print("Pontos: ", linha_points)
# #                 print("Pergunta: ", pergunta)
# #                 print("Linha: ", linha)
# #                 print("NGram_Tupla: ", ngram_tupla_texto)
#         if linha_points > 3:
#             print("Linha: ", linha)
#             print("Corresponde a pergunta: ", pergunta.lower())
```