

✖ Importação das bibliotecas

```
O código gerado pode estar sujeito a uma licença | 03Akshay/assignments-1 | AEdgecombe/ML-Testing
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
```




✖ Ciando a base de dados

```
data = {
    'email': [
        # Spam
        'Ganhe dinheiro fácil trabalhando em casa',
        'Oferta imperdível! Compre agora',
        'Parabéns! Você ganhou um prêmio',
        'Clique aqui para ganhar um iPhone',
        'Última chance! Descontos incríveis',
        'Você foi selecionado para um prêmio exclusivo',
        'Ganhe até R$5000 por semana sem sair de casa',
        'Seu CPF foi premiado, clique aqui para confirmar',
        'Oferta exclusiva! Só hoje, compre com 70% de desconto',
        'Emprego dos sonhos, sem sair de casa, ganhos ilimitados',
        'Promoção imperdível, compre dois e leve três',
        'Confirme seus dados e receba seu prêmio imediatamente',
        'Você tem uma dívida pendente, clique para negociar',
        'Seja seu próprio chefe, comece hoje mesmo',
        'Aumente seu limite de crédito sem burocracia',
        'Aproveite! Últimas unidades com preço de fábrica',
        'Liberamos um crédito especial para seu CPF',
        'Ganhe seguidores e curtidas no Instagram instantaneamente',
        'Seu empréstimo foi aprovado, clique para sacar agora',
        'Você foi sorteado! Retire seu prêmio agora mesmo',

        # Não spam
        'Sua fatura está disponível',
        'Reunião agendada para amanhã',
        'Segue o relatório solicitado',
        'Pedido confirmado com sucesso',
        'Vamos almoçar hoje?',
        'Envio da nota fiscal do seu pedido',
        'Atualização sobre seu pedido de suporte',
        'Lembrete: reunião às 15h de amanhã',
        'Relatório financeiro do mês de junho',
        'Confirmação de participação no evento',
        'Análise de desempenho do último trimestre',
        'Documentação necessária para atualização cadastral',
        'Parabéns pelo seu aniversário, felicidades!',
        'A sua assinatura foi renovada com sucesso',
        'Informações sobre a manutenção programada',
        'Sua senha foi alterada com sucesso',
        'Bem-vindo à nossa plataforma, aproveite os benefícios',
        'Reembolso processado, valor estará disponível em 3 dias',
        'Agendamento da consulta confirmado para sexta-feira',
        'Atualização dos termos de serviço da plataforma'
    ],
    'label': [
        1, 1, 1, 1, 1,
        1, 1, 1, 1, 1,
        1, 1, 1, 1, 1,
        1, 1, 1, 1, 1,
        0, 0, 0, 0, 0,
        0, 0, 0, 0, 0,
        0, 0, 0, 0, 0,
        0, 0, 0, 0, 0
    ]
}
```

```
df = pd.DataFrame(data)
df
```



	email	label	
0	Ganhe dinheiro fácil trabalhando em casa	1	
1	Oferta imperdível! Compre agora	1	
2	Parabéns! Você ganhou um prêmio	1	
3	Clique aqui para ganhar um iPhone	1	
4	Última chance! Descontos incríveis	1	
5	Você foi selecionado para um prêmio exclusivo	1	
6	Ganhe até R\$5000 por semana sem sair de casa	1	
7	Seu CPF foi premiado, clique aqui para confirmar	1	
8	Oferta exclusiva! Só hoje, compre com 70% de d...	1	
9	Emprego dos sonhos, sem sair de casa, ganhos i...	1	
10	Promoção imperdível, compre dois e leve três	1	
11	Confirme seus dados e receba seu prêmio imedia...	1	
12	Você tem uma dívida pendente, clique para nego...	1	
13	Seja seu próprio chefe, comece hoje mesmo	1	
14	Aumente seu limite de crédito sem burocracia	1	
15	Aproveite! Últimas unidades com preço de fábrica	1	
16	Liberamos um crédito especial para seu CPF	1	
17	Ganhe seguidores e curtidas no Instagram insta...	1	
18	Seu empréstimo foi aprovado, clique para sacar...	1	
19	Você foi sorteado! Retire seu prêmio agora mesmo	1	
20	Sua fatura está disponível	0	
21	Reunião agendada para amanhã	0	
22	Segue o relatório solicitado	0	
23	Pedido confirmado com sucesso	0	
24	Vamos almoçar hoje?	0	
25	Envio da nota fiscal do seu pedido	0	
26	Atualização sobre seu pedido de suporte	0	
27	Lembrete: reunião às 15h de amanhã	0	
28	Relatório financeiro do mês de junho	0	
29	Confirmação de participação no evento	0	
30	Análise de desempenho do último trimestre	0	
31	Documentação necessária para atualização cadas...	0	
32	Parabéns pelo seu aniversário, felicidades!	0	
33	A sua assinatura foi renovada com sucesso	0	
34	Informações sobre a manutenção programada	0	
35	Sua senha foi alterada com sucesso	0	
36	Bem-vindo à nossa plataforma, aproveite os ben...	0	
37	Reembolso processado, valor estará disponível ...	0	
38	Agendamento da consulta confirmado para sexta-...	0	
39	Atualização dos termos de serviço da plataforma	0	

Próximas etapas:

[Gerar código com df](#)

[Ver gráficos recomendados](#)

[New interactive sheet](#)

✖ Vetorização dos textos

```
vectorizer = CountVectorizer()
X = vectorizer.fit_transform(df['email'])
y = df['label']
```

✖ Divisão dos dados

```
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

✖ Treinamento (Regressão logística)

```
modelo_reg = LogisticRegression()
modelo_reg.fit(x_train, y_train)
```

LogisticRegression

LogisticRegression()

Avaliação

```
y_pred = modelo_reg.predict(x_test)
print("Acurácia:", accuracy_score(y_test, y_pred))
print("\nMatriz de Confusão:", confusion_matrix(y_test, y_pred))
print("\nRelatório de Classificação:", classification_report(y_test, y_pred))
```

Acurácia: 0.75

Matriz de Confusão: [[3 0]
[2 3]]

Relatório de Classificação:

				precision	recall	f1-score	support
	0	0.60	1.00	0.75	3		
	1	1.00	0.60	0.75	5		
	accuracy			0.75	8		
	macro avg	0.80	0.80	0.75	8		
	weighted avg	0.85	0.75	0.75	8		

Testes

```
while True:
    novo_email = input("\nDigite o texto do e-mail para classificar (ou 'sair' para encerrar):\n")
    if novo_email.lower() == 'sair':
        break
    novo_email_vect = vectorizer.transform([novo_email])
    pred = modelo_reg.predict(novo_email_vect)

    if pred[0] == 1:
        print("🔴 Este e-mail é classificado como: SPAM\n")
    else:
        print("🟢 Este e-mail é classificado como: NÃO SPAM\n")
```

Digite o texto do e-mail para classificar (ou 'sair' para encerrar):
sair

Naive Bayes

```
modelo_naive_bayes = MultinomialNB()
modelo_naive_bayes.fit(x_train, y_train)
```

MultinomialNB

MultinomialNB()

```
#Predição
y_pred_naive_bayes = modelo_naive_bayes.predict(x_test)
#Avaliação
print("Acurácia:", accuracy_score(y_test, y_pred_naive_bayes))
print("\nMatriz de Confusão:", confusion_matrix(y_test, y_pred_naive_bayes))
print("\nRelatório de Classificação:", classification_report(y_test, y_pred_naive_bayes))
```

Acurácia: 0.75

Matriz de Confusão: [[3 0]
[2 3]]

Relatório de Classificação:

				precision	recall	f1-score	support
	0	0.60	1.00	0.75	3		
	1	1.00	0.60	0.75	5		
	accuracy			0.75	8		
	macro avg	0.80	0.80	0.75	8		
	weighted avg	0.85	0.75	0.75	8		