

Classificação de Sentimentos de LinearSVC para Loja Online

Alunos:

Pedro Henrique Veras – 2120597

André Pereira dos Santos – 2120481

Pedro Afonso Ferreira Leite – 2120431

José Augusto Pires de Aleluia - 2120605

Contextualização:

A análise de sentimentos é um campo da inteligência artificial que busca extrair e classificar emoções, opiniões e atitudes de textos. Essa tecnologia tem se tornado cada vez mais popular e relevante em diversos setores, especialmente no comércio online.

Justificativa:

A classificação de sentimentos em lojas online apresenta diversas vantagens:

1. Melhoria da experiência do cliente:

- Identificar e solucionar problemas com produtos ou serviços de forma mais rápida e eficiente.
- Personalizar a comunicação com os clientes, oferecendo ofertas e promoções relevantes.
- Obter feedback valioso para aprimorar a experiência do cliente.

2. Aumento das vendas:

- Identificar produtos com alta demanda e melhorar sua visibilidade.
- Oferecer produtos e serviços personalizados para cada cliente.
- Aumentar a taxa de conversão de vendas.

3. Fortalecimento da marca:

- Monitorar a reputação da marca online.
- Identificar e responder a comentários negativos de forma eficaz.
- Construir uma imagem positiva da marca junto aos clientes.

Relevância no Contexto Atual:

Com o crescimento do comércio online, a análise de sentimentos se torna ainda mais importante para as empresas se destacarem no mercado. A capacidade de entender e responder às necessidades dos clientes de forma personalizada é fundamental para o sucesso de qualquer loja online.

Objetivo Geral:

Implementar um sistema de classificação de sentimentos em uma loja online para analisar e entender as emoções, opiniões e atitudes dos clientes em relação à loja, seus produtos e serviços.

Objetivos Específicos:

1. Coletar e preparar dados de treinamento:

- Coletar um conjunto de dados de texto com rótulos de sentimento (positivo, negativo, neutro) de diversas fontes:
 - Comentários e avaliações de produtos na loja online.
 - Feedback dos clientes através de formulários de contato, chat online, etc.
 - Redes sociais da loja (Facebook, Instagram, Twitter, etc.).
- Dividir o conjunto de dados em treino e teste.

2. Treinar e avaliar o modelo de classificação de sentimentos:

- Treinar um modelo LinearSVC com o conjunto de dados de treino.
- Avaliar a acurácia do modelo no conjunto de dados de teste.
- Ajustar os parâmetros do modelo (C, penalty, etc.) para otimizar a acurácia.

3. Implementar o modelo na loja online:

- Salvar o modelo treinado em um arquivo.
- Criar um script Python para carregar o modelo e classificar novos textos.

4. Analisar os resultados e tomar decisões:

- Identificar áreas de melhoria na loja online com base nos sentimentos dos clientes.
- Tomar decisões estratégicas para melhorar a experiência do cliente, aumentar as vendas e fortalecer a marca.

A metodologia utilizada para o projeto de classificação de sentimentos em loja online pode ser dividida em cinco etapas principais:

1. Coleta e preparação de dados:

• Fontes de dados:

- Comentários e avaliações de produtos na loja online.
- Feedback dos clientes através de formulários de contato, chat online, etc.
- Redes sociais da loja.

• Pré-processamento de dados:

- Remoção de stopwords, pontuação, caracteres especiais, etc.
- Normalização do texto (minúsculas, lematização, etc.).

- Divisão do conjunto de dados em treino e teste.

2. Treinamento do modelo:

- **Algoritmo:** LinearSVC.
- **Parâmetros:** C, penalty, max_iter, etc.
- **Métricas de avaliação:** acurácia, precisão, rechadação, F1-score.

3. Implementação do modelo:

- **Salvar o modelo treinado:** arquivo pickle.
- **Classificação de novos textos:** script Python.

4. Análise dos resultados:

- **Identificação de áreas de melhoria:** produtos, serviços, experiência do cliente.
- **Tomada de decisões estratégicas:** aumento de vendas, fidelização de clientes, fortalecimento da marca.

Resultados e Avaliação do Modelo

# Exemplo de dados de treinamento				
data = {				
'texto': ["Este produto é incrível!", "Não gostei muito deste produto.", "O produto não atendeu às minhas expectativas.", "Uma belíssima porcaria!"],				
'sentimento': ['positivo', 'negativo', 'negativo', 'negativo']				
}				
TERMINAL				
	precision	recall	f1-score	support
negativo	1.00	1.00	1.00	1
accuracy			1.00	1
macro avg	1.00	1.00	1.00	1
weighted avg	1.00	1.00	1.00	1

1. **Divisão de Dados:** O código divide os dados em conjuntos de treinamento e teste usando train_test_split. 80% dos dados são usados para treinamento e 20% para teste.
2. **Vetorização de Texto:** Os dados de texto são vetorizados usando TfidfVectorizer, que converte o texto em uma representação numérica usando o esquema TF-IDF (Term Frequency-Inverse Document Frequency).
3. **Treinamento do Modelo:** O modelo LinearSVC é treinado com os dados de treinamento vetorizados.
4. **Avaliação do Modelo:** O modelo treinado é avaliado usando os dados de teste. Ele calcula a precisão, recall, F1-score e outras métricas usando classification_report.
5. **Classificação de Nova Avaliação:** Uma nova avaliação é fornecida para o modelo, e ele prediz o sentimento associado a essa avaliação.

Avaliando o resultado:

- O desempenho do modelo pode ser avaliado observando as métricas do classification_report, como precisão, recall e F1-score. Essas métricas fornecem uma ideia de quão bem o modelo está generalizando para dados não vistos.

- A predição do sentimento para a nova avaliação pode ser verificada manualmente para ver se corresponde às expectativas.