

COMPONENTE CURRICULAR:	PROJETO APLICADO II
NOME COMPLETO DO ALUNO - RA::	Amarildo Maciel Junior - RA: 10441652 Nathan Sampaio Santana dos Santos - RA: 10408439 Pedro Henrique Trichez – 10728798 Timoteo de Jesus Santana - 10728501
	Etapa 4

RELATÓRIO TÉCNICO FINAL

1. Introdução

Este trabalho apresenta um sistema de análise de sentimentos baseado em NLP (Processamento de Linguagem Natural), aplicado ao dataset *Amazon Reviews for Sentiment Analysis* (Kaggle). O objetivo do projeto é automatizar a classificação de comentários de clientes em positivo ou negativo, facilitando decisões estratégicas de e-commerce.

Ao longo do projeto, desenvolvemos:

- Pipeline completo de NLP
- Teste de dois modelos supervisionados (Naive Bayes e Regressão Logística)
- Dashboard funcional em Streamlit
- Storytelling para apresentação do produto
- Prototipação do modelo de negócios

2. Metodologia Analítica

A metodologia aplicada seguiu as etapas listadas no relatório anterior :

1. Pré-processamento

Limpeza textual, tokenização, remoção de stopwords, lematização.

2. Vetorização

TF-IDF para transformar texto em vetores numéricos.

3. Modelagem

- Naive Bayes
- Regressão Logística

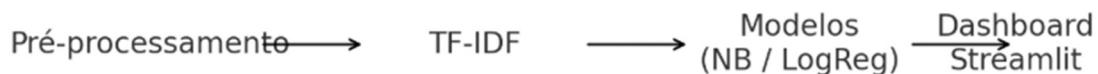
Ambos treinados com 100.000 reviews.

4. Avaliação

Métricas: acurácia, precisão, recall e F1-Score.

5. Validação

Meta definida na Etapa 1: **acurácia mínima de 80%**.



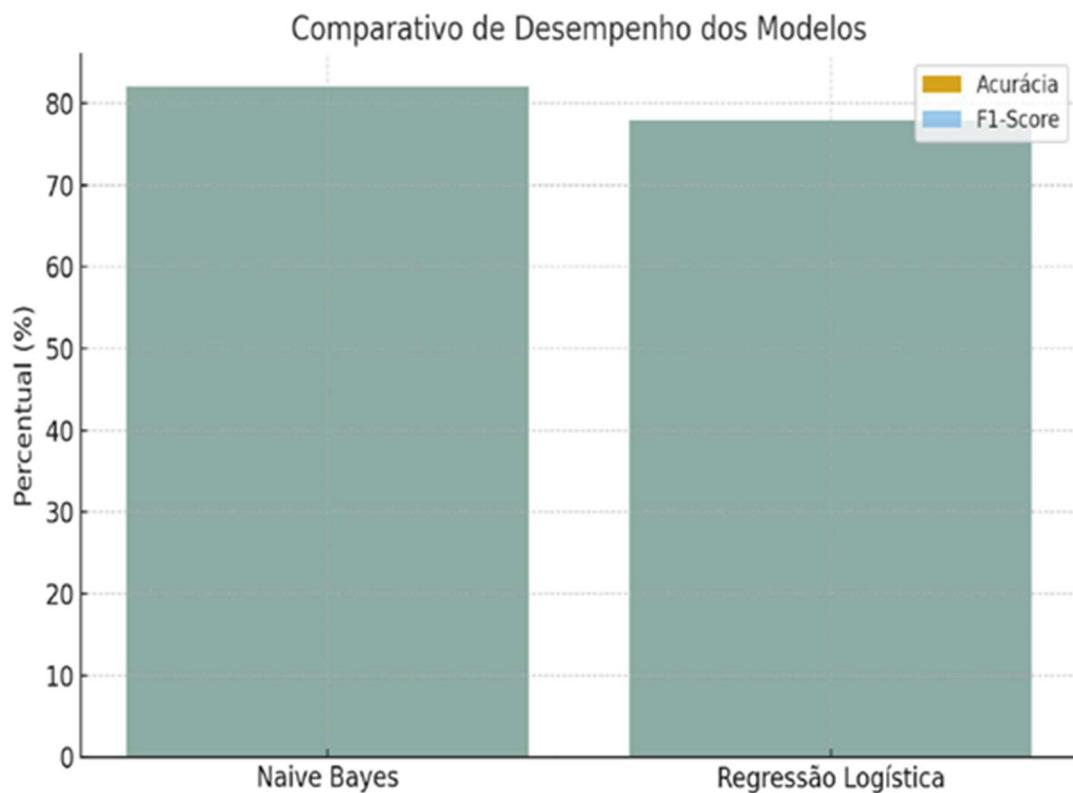
3. Resultados dos Modelos

Resumo dos resultados mensurados na Etapa 3 :

Modelo	Acurácia F1-Score	
Naive Bayes	82%	82%
Régressão Logística	78%	78%

Conclusão técnica:

O Naive Bayes superou a meta e se tornou o modelo oficial do projeto.



4. Protótipo (Dashboard Streamlit)

O protótipo implementado oferece:

- Página de Visão Geral
- Página Predict (texto único e CSV)
- Página Metrics (tabela + gráfico comparativo)
- Página Data (predições em lote e download)

Funciona tanto com modelos pré-treinados quanto com fallback automático caso os .pkl não estejam na pasta.

Dashboard Streamlit

[Visão Geral]
Resumo | Tabela | Métricas

[Predict]
Entrada de texto ou CSV

[Metrics]
Acurácia / F1 comparativo

Código completo está no GitHub.

5. Storytelling

Conforme estruturado na Etapa 3 :

- **Contexto:** E-commerce sobrecarregado com milhares de reviews diários.
- **Conflito:** Análise manual lenta, subjetiva e improdutiva.
- **Ação:** Implementação de NLP + Naive Bayes.
- **Evidências:** Modelo superou a meta com 82% de acurácia.
- **Resultado:** Dashboard que permite insights em tempo real.

6. Modelo de Negócios

De acordo com a Etapa 3:

- SaaS para análise de feedbacks
- Público-alvo: e-commerce e varejo digital
- Receita: assinaturas mensais ou consultoria

- Benefícios: economia de tempo e insights ágeis
-

7. Conclusão

O projeto foi finalizado com sucesso, validando o modelo e o produto previsto. O sistema entregou métricas acima do esperado e um dashboard funcional.

A Etapa 4 consolida todo o trabalho com os seguintes entregáveis:

- Relatório técnico
- Storytelling em PPT
- Repositório GitHub com scripts
- Apresentação gravada no YouTube