

UNIVERSIDADE PRESBITERIANA MACKENZIE

BRENNO MONTEIRO DE OLIVEIRA

BRUNO FERNANDES MASCARINI

LUCAS CAMARGO SPINELI

MARINA CAMARGO SPINELI

PROJETO APLICADO II

ANÁLISE DAS PERCEPÇÕES DO CONTEÚDO DO TWITTER

São Paulo
2025

BRENNO MONTEIRO DE OLIVEIRA
BRUNO FERNANDES MASCARINI
LUCAS CAMARGO SPINELI
MARINA CAMARGO SPINELI

PROJETO APLICADO II

ANÁLISE DAS PERCEPÇÕES DO CONTEÚDO DO TWITTER

Projeto Aplicado II, do terceiro período do curso
de Tecnologia em Ciência de Dados da
Universidade Presbiteriana Mackenzie.

Prof. Dr. Felipe Albino dos Santos

São Paulo
2025

SUMÁRIO

1. INTRODUÇÃO.....4

2. CRONOGRAMA.....5

3. OBJETIVOS DO ESTUDO.....6

4. CONTEXTO DE ESTUDO.....7

5. APRESENTAÇÃO DA EMPRESA.....8

6. APRESENTAÇÃO DO DATASET.....9

1. INTRODUÇÃO

O suporte ao cliente nas redes sociais tem se tornado uma das principais formas de interação entre empresas e consumidores, e o X (Twitter), com sua natureza dinâmica e interativa, se destaca como uma das plataformas mais utilizadas para esse fim. Com mais de 3 milhões de tweets e respostas de grandes marcas, o conjunto de dados analisado neste estudo oferece uma oportunidade de explorar a comunicação entre empresas e consumidores, além de fornecer insights sobre a eficácia das práticas de suporte ao cliente modernas.

Este trabalho se propõe a investigar esses dados, visando compreender melhor as interações em tempo real, a análise da percepção do consumidor e o impacto das respostas oferecidas pelas empresas. A análise desses dados é relevante não apenas para aprimorar as estratégias de atendimento ao cliente, mas também para contribuir com a evolução da compreensão da linguagem natural e o desenvolvimento de modelos conversacionais. Ao longo deste estudo, exploraremos as características do conjunto de dados, suas vantagens em relação a outros corpora conversacionais e as possibilidades de aplicação, incluindo a previsão de respostas, o agrupamento de tópicos e a identificação de novos problemas que possam surgir no contexto do atendimento online.

Toda a documentação, ficará disponível no github, no link a seguir:

<https://github.com/Brennu/Projeto-Aplicado-II>

2. CRONOGRAMA

O cronograma a seguir ilustra as etapas a serem realizadas neste projeto. Cada integrante do grupo está representado por uma cor, indicando as respectivas responsabilidades de cada um; o “X” marca a expectativa da semana do mês de realização de cada parte do trabalho.

LEGENDA					
	BRENNO MONTEIRO DE OLIVEIRA		LUCAS CAMARGO SPINELI		TODOS
	BRUNO FERNANDES MASCARINI		MARINA CAMARGO SPINELI		

CRONOGRAMA - PROJETO APLICADO II														
	Lista de Atividades	RESP.	FEVEREIRO			MARÇO			ABRIL			MAIO		
Etapa 1	Definição da Organização		X											
	Apresentação do Dados (Metadados)		X											
	Objetivos e Metas		X											
	Cronograma de Atividades			X										
	Estrutura do Documento		X	X										
	Link para o Github		X											
Etapa 2	Definição da Linguagem de Programação					X								
	Análise Exploratória da Base de Dados					X	X							
	Tratamento da Base de Dados					X								
	Definição e Descrição das Bases Teóricas dos Métodos							X						
	Definição e Descrição da Acurácia							X						
Etapa 3	Aplicação do Método Analítico							X						
	Medidas de Acurácia							X						
	Descrição dos Resultados Preliminares							X						
	Esboço do Storytelling								X					
Etapa 4	Relatório Técnico										X			
	Apresentação do Storytelling											X	X	
	Repositório no GitHub												X	
	Vídeo da Apresentação												X	X

3. OBJETIVOS DO ESTUDO

O presente estudo tem como objetivo desenvolver um modelo capaz de analisar e classificar sentimentos em tweets, identificando padrões e tendências nas opiniões expressas pelos usuários. Para isso, será realizada uma investigação detalhada dos dados, incluindo a limpeza e o pré-processamento dos textos, a aplicação de algoritmos de aprendizado de máquina e a avaliação do desempenho dos modelos. Ao final, pretende-se construir uma ferramenta eficiente para prever a polaridade dos sentimentos em novas postagens no Twitter, contribuindo para uma melhor compreensão do comportamento dos usuários na plataforma.

4. CONTEXTO DE ESTUDO

Com o crescimento das redes sociais, o Twitter tornou-se uma das principais plataformas onde os usuários compartilham suas opiniões sobre uma ampla variedade de assuntos, como produtos, serviços, eventos e figuras públicas. A grande quantidade de postagens diárias cria um ambiente rico em informações, mas ao mesmo tempo desafiador para análise manual. Diante desse cenário, a análise automatizada de sentimentos surge como uma ferramenta essencial para compreender as percepções dos usuários de forma rápida e eficiente.

Empresas e organizações utilizam a análise de sentimentos para monitorar a recepção de seus produtos, avaliar a satisfação do público e até mesmo antecipar crises de imagem. Ao identificar padrões nas emoções expressas nos tweets, é possível obter insights valiosos que auxiliam na tomada de decisões estratégicas.

Além disso, com o avanço do processamento de linguagem natural e aprendizado de máquina, é possível identificar nuances emocionais complexas, como sarcasmo e ironia. Isso torna a análise de sentimentos mais precisa e permite que empresas respondam rapidamente a problemas de clientes no Twitter. Assim, a análise de sentimentos se torna uma ferramenta estratégica, ajudando as organizações a adaptar suas estratégias e melhorar a experiência do cliente.

5. APRESENTAÇÃO DA EMPRESA

O X, anteriormente conhecido como Twitter, é uma plataforma de mídia social dedicada à comunicação em tempo real. Fundada em 2006 por Jack Dorsey, Biz Stone e Evan Williams, a empresa tem sede em São Francisco, Califórnia, EUA.

Missão: A missão do X é promover e proteger a conversa pública — ser a praça pública da internet.

Visão: Ser o melhor aplicativo para tudo que gira em torno de áudio, vídeo, mensagens, pagamentos e serviços bancários

Valores: Os valores do X refletem seu compromisso com a comunicação aberta e inclusiva. A empresa incentiva a comunicação sem medo para construir confiança, garantindo um ambiente onde os usuários possam se expressar livremente. Além disso, busca defender e respeitar a voz do usuário, promovendo um espaço de diálogo autêntico.

Segmento de atuação: O X opera no setor de tecnologia e redes sociais, oferecendo um espaço para debates, notícias, entretenimento e interação entre usuários, empresas e governos. Seu modelo de negócios inclui publicidade digital, assinaturas e monetização de conteúdo, atendendo a um público global.

Market Share: O X sofreu uma desvalorização de aproximadamente 79% desde a aquisição por Elon Musk, reduzindo seu valor de mercado para cerca de US\$ 9,4 bilhões. Relatórios indicam que o número de usuários ativos mensais do X caiu para aproximadamente 335 milhões em 2024, uma redução em relação aos 368 milhões anteriores.

Iniciativas em Ciência de Dados: O X utiliza técnicas de ciência de dados para analisar sentimentos e tendências nas postagens dos usuários, auxiliando empresas a entenderem a percepção de suas marcas e produtos.

6. APRESENTAÇÃO DO DATASET

O dataset utilizado neste estudo "Twitter Sentiment Analysis" contém informações extraídas do Twitter e apresenta quatro tipos de dados principais. Cada entrada no conjunto de dados possui um identificador único que diferencia os tweets entre si. Além disso, há uma coluna que representa a categoria ou tópico principal abordado na postagem, como o nome de um produto, jogo ou assunto específico discutido pelos usuários.

Outro dado relevante presente no dataset é a classificação do sentimento expresso no tweet. Cada postagem é rotulada de acordo com a emoção predominante, podendo ser positiva, negativa, neutra ou irrelevante. Por fim, o conjunto de dados inclui o próprio texto do tweet, que contém a opinião do usuário sobre o tema abordado. Esses dados permitirão a aplicação de técnicas de Processamento de Linguagem Natural para compreender e classificar automaticamente os sentimentos expressos nas postagens.