

FATEC DE REGISTRO
DESENVOLVIMENTO DE SOFTWARE MULTIPLATAFORMA

AMANDA NOGUEIRA DE CASTRO E SILVA

ADELDIVO ALVES DE SOUSA JUNIOR

ESTUDO DE CASO:
ANÁLISE DE SENTIMENTO DE AVALIAÇÕES DE PRODUTOS

ESTUDO DE CASO: ANÁLISE DE SENTIMENTO DE AVALIAÇÕES DE PRODUTOS

Amanda Nogueira de Castro e Silva, Adeldivo Alves de Sousa Junior

RESUMO:

Este estudo de caso tem como objetivo aplicar conceitos de Processamento de Linguagem Natural (PLN) para automatizar a análise de sentimentos de avaliações de produtos em um e-commerce. Devido ao grande volume de feedbacks recebidos diariamente, torna-se inviável a análise manual. A solução proposta classifica automaticamente as avaliações dos clientes em positivas, negativas ou neutras, além de identificar palavras-chave e temas principais mencionados nas avaliações. Este relatório descreve a metodologia adotada para coleta, pré-processamento dos dados, e criação de uma base de dados para treinamento e teste, visando melhorar a compreensão do feedback dos clientes e apoiar decisões estratégicas da empresa.

PALAVRAS-CHAVE: Processamento de Linguagem Natural. Análise de Sentimentos. Avaliação de Produtos.

1. INTRODUÇÃO:

Empresas de *e-commerce* enfrentam o desafio de entender e analisar o *feedback* de seus clientes, que frequentemente deixam avaliações e comentários sobre suas compras. Com milhares de avaliações postadas diariamente, a equipe de suporte ao cliente não consegue ler e analisar manualmente todas elas, o que dificulta a identificação de problemas com produtos específicos, a melhora da satisfação do cliente, e a tomada de decisões mais informadas sobre quais produtos promover ou descontinuar.

Para solucionar esse problema, este estudo de caso propõe o uso de Processamento de Linguagem Natural (PLN) para automatizar a análise de sentimentos das avaliações dos clientes. O objetivo principal é classificar automaticamente as avaliações como positivas, negativas ou neutras e identificar as principais palavras-chave e temas mencionados, proporcionando insights valiosos para a empresa.

2. METODOLOGIA

2.1- Coleta de Dados

As avaliações foram coletadas do site Mercado Livre do produto *Samsung Galaxy A15 Dual SIM 4G 256GB Azul escuro 8GB RAM*. Foram selecionadas 21 avaliações que já estavam rotuladas com base nas classificações fornecidas pelos clientes (1 a 5 estrelas). Avaliações com 1-2 estrelas foram classificadas como negativas, 3 estrelas como neutras, e 4-5 estrelas como positivas.

2.2- Pré-processamento de Dados

1. Limpeza de Dados:

- Remoção de ruídos como pontuação, números, caracteres especiais e palavras irrelevantes (*stopwords*).
- A remoção de palavras irrelevantes incluiu artigos, preposições, conjunções, pronomes e outras palavras de uso frequente que não contribuem para o entendimento do contexto.

2. Tokenização:

- Cada avaliação foi dividida em palavras (*tokens*). Por exemplo, a frase "O produto é excelente" se tornou ["produto", "excelente"].

2.3- Criação da Base de Dados

1. Classificação das Avaliações:

- As avaliações foram classificadas em três categorias: positivas, negativas e neutras.

2. Divisão da Base de Dados:

- A base de dados foi dividida em duas partes: treinamento e teste, utilizando a técnica de *holdout* estratificado. Essa técnica garante que ambas as partes mantenham a proporção de avaliações positivas, negativas e neutras, garantindo um treinamento equilibrado para o modelo.

3. RESULTADOS

A seguir, apresentamos os resultados obtidos a partir do processamento das avaliações:

- **Dados Processados:** As avaliações foram limpas e *tokenizadas*, removendo elementos que não contribuem para a análise de sentimento.
- **Divisão de Dados:** A base de dados foi dividida em conjuntos de treinamento (80%) e teste (20%) de maneira estratificada, garantindo representatividade proporcional de cada classe.

Exemplo de uma avaliação *tokenizada* e classificada:

- Avaliação Original: "O produto é excelente, recomendo!"
- Avaliação Limpa: "produto excelente recomendo"
- Tokenização: ['produto', 'excelente', 'recomendo']
- Classificação: Positiva

4. DISCUSSÃO

A análise automatizada de sentimentos utilizando PLN demonstrou ser eficaz para entender rapidamente o *feedback* dos clientes. A técnica de pré-processamento de dados permitiu a remoção de ruídos, melhorando a qualidade dos dados de entrada para o modelo de análise de sentimento. A tokenização ajudou a identificar palavras-chave que são críticas para entender o sentimento geral das avaliações.

Possíveis melhorias incluem a implementação de modelos de *machine learning* mais avançados, como redes neurais recorrentes (RNN¹) ou *transformers*², que poderiam capturar melhor o contexto e a nuance do sentimento nas avaliações dos clientes. Além disso, seria interessante explorar técnicas de análise de sentimentos mais refinadas, como a detecção de sarcasmo, que pode afetar a precisão da classificação.

5. CONCLUSÃO

O uso de Processamento de Linguagem Natural para automatizar a análise de sentimentos das avaliações dos clientes mostrou-se uma abordagem promissora para melhorar a compreensão do *feedback* do cliente e apoiar decisões estratégicas em empresas de *e-commerce*. A solução proposta permite identificar rapidamente problemas com produtos específicos e tomar decisões informadas sobre quais produtos promover ou descontinuar, melhorando a satisfação do cliente e o desempenho geral da empresa.

6. APÊNDICES

- **Apêndice A:** Tabela de Limpeza das Avaliações
- **Apêndice B:** Tabela de Tokenização das Avaliações
- **Apêndice C:** Database de Treinamento
- **Apêndice D:** Database de Teste

¹ **RNN:** É um tipo de modelo de aprendizado profundo projetado para processar e transformar uma sequência de entrada de dados em uma sequência de saída de dados específica (AWS, 2023).

² **Transformes:** Os *transformers* são uma arquitetura de aprendizado profundo projetada para resolver tarefas *sequence-to-sequence*, ao mesmo tempo em que lida facilmente com dependências de longo alcance (DSA, 2022).

7. REFERÊNCIAS:

MERCADO LIVRE. **Samsung Galaxy A15 Dual SIM 4G 256GB Azul escuro 8GB RAM**. Disponível em: <https://www.mercadolivre.com.br/samsung-galaxy-a15-dual-sim-4g-256gb-azul-escuro-8gb-ram/p/MLB31214368>. Acesso em: 1 set. 2024.

AWS. **O que é RNN (Rede neural recorrente)**. 2023. Disponível em: [https://aws.amazon.com/pt/what-is/recurrent-neural-network/#:~:text=Uma%20Recurrent%20Neural%20Network%20\(RNN,sa%C3%ADda%20de%20dados%20sequencial%20espec%C3%ADfica](https://aws.amazon.com/pt/what-is/recurrent-neural-network/#:~:text=Uma%20Recurrent%20Neural%20Network%20(RNN,sa%C3%ADda%20de%20dados%20sequencial%20espec%C3%ADfica). Acesso em: 1 set. 2024.

DSA, E. **Capítulo 85 - Transformadores - O Estado da Arte em Processamento de Linguagem Natural**. 2022. Disponível em: <https://www.deeplearningbook.com.br/transformadores-o-estado-da-arte-em-processamento-de-linguagem-natural/>. Acesso em: 1 set. 2024.

APÊNDICE A – TABELA DE LIMPEZA DAS AVALIAÇÕES

Avaliação Original	Avaliação Limpa
"É amplamente elogiado por sua rapidez, ótima câmera e bateria durável, além de não travar durante o uso. O aparelho também é apreciado por suas funções adicionais como NFC e Smart View, e pelo excelente custo-benefício."	"amplamente elogiado rapidez ótima câmera bateria durável não travar durante uso apreciado funções nfc smart view custo-benefício"
"Gostei bastante, a tela tem um ótimo tamanho... achei bem satisfatório para o preço do celular, não tenho com o que reclamar por enquanto, estou bem satisfeito!"	"gostei bastante tela ótimo tamanho satisfatório preço celular não tenho reclamar satisfeito"
"Celular muito bonito, veio tudo certinho!! só não gostei pq não embalaram o cell em um saco bolha, e a caixa veio sem nenhuma proteção! mas o telefone não tem nenhum arranhão. Tirando isso, tudo certo."	"celular bonito tudo certo não gostei não embalaram telefone saco bolha caixa sem proteção telefone sem arranhão tudo certo"
"Dei de presente pra minha esposa... Por enquanto tudo certo."	"dei presente esposa tudo certo"
"Por enquanto estou achando muito bom, fotos boas, respostas rápidas da biometria, além disso tem o smart view... fora isso ele é muito bom. Obs o carregador é fraco demorar para carregar o telefone um pouco."	"achando muito bom fotos boas respostas rápidas biometria smart view bom carregador fraco demora carregar telefone"
"Produto muito bom bateria dura bastante só não gostei do tamanho, estou acostumado com celular maior, mas pra quem gosta recomendo podem comprar sem medo."	"produto muito bom bateria dura bastante não gostei tamanho acostumado celular maior recomendo comprar sem medo"
"O produto tem um design lindo, porém tem pouca durabilidade em relação à bateria... Só não faço a devolução devido a necessidade que eu tenho em relação ao aparelho."	"produto design lindo pouca durabilidade bateria não devolvo necessidade aparelho"
"Eu não tive muito sorte com essa compra, esse celular está com um mês de uso e tá travando... ele não obedece."	"não tive sorte compra celular mês uso travando não obedece"
"A qualidade das câmeras deixa muito a desejar, mas para um básico, dá pro gasto... precisa adquirir um separadamente."	"qualidade câmeras deixa desejar básico dá gasto precisa adquirir separadamente"
"Celular mediano, esperava mais por ser samsung... travando o app usado no momento fazendo com que você desigualdade tela para reiniciar o app."	"celular mediano esperava mais samsung travando app momento desigualdade tela reiniciar app"
"Produto infelizmente após poucos meses de uso, apresenta defeitos de travamento e acessos remotos a apps e ligações."	"produto após meses uso defeitos travamento acessos remotos apps ligações"
"Produto muito bom bateria dura bastante só não gostei do tamanho... mas pra quem gosta recomendo podem comprar sem medo."	"produto muito bom bateria dura bastante não gostei tamanho recomendo comprar sem medo"
"O produto tem um design lindo, porém tem pouca durabilidade em relação à bateria... Só não faço a devolução devido a necessidade que eu tenho em relação ao aparelho."	"produto design lindo pouca durabilidade bateria não devolvo necessidade aparelho"
"Ou configuração pois não me adaptei com o a15..."	"configuração não adaptei a15 faz parte não ruim"

mas faz parte, não é ruim mas a configuração tem que sofrer pois já fui em 2 técnico samsung."	configuração sofrer já fui técnico samsung"
"Produto veio errado e me foi negado o reembolso de forma descarada e ilegal com total desprezo ao consumidor."	"produto errado negado reembolso descarada ilegal desprezo consumidor"
"O meu só prestou 3 dias aí não carregou mais e eles não aceitaram a devolução fiquei no prejuízo."	"meu prestou 3 dias não carregou mais não aceitaram devolução fiquei prejuízo"
"Péssima qualidade. Comprei com 1 mes de uso deu problema não tá carregando mais. Péssimo não comprem."	"péssima qualidade 1 mês uso problema não carregando péssimo não comprem"
"É um celular bem lento. Bem devagar para abrir apps e executar tarefas simples como copiar e colar texto entre dois apps."	"celular bem lento devagar abrir apps executar tarefas simples copiar colar texto dois apps"
"O celular com 1 mês de uso apresentou problemas... Vou ter q mandar pra assistência técnica."	"celular 1 mês uso apresentou problemas mandar assistência técnica"
"O celular é bom, por onde sai o áudio que parece que não tem uma proteção é como se fosse outro celular que encaixaram em cima dele só tem metade da proteção do audio na parte de cima."	"celular bom áudio parece não proteção encaixaram cima metade proteção áudio cima"
"Nem cinco dia de uso já começou a dar defeito... Mas o celular não 🚫 🗑️ . Minha primeira experiência, foi a desejar."	"cinco dias uso defeito celular não primeira experiência desejar"

APÊNDICE B – TABELA DE TOKENIZAÇÃO DAS AVALIAÇÕES

Avaliação Limpa	Tokenização
"amplamente elogiado rapidez ótima câmera bateria durável não travar durante uso apreciado funções nfc smart view custo-benefício"	['amplamente', 'elogiado', 'rapidez', 'ótima', 'câmera', 'bateria', 'durável', 'não', 'travar', 'durante', 'uso', 'apreciado', 'funções', 'nfc', 'smart', 'view', 'custo-benefício']
"gostei bastante tela ótimo tamanho satisfatório preço celular não tenho reclamar satisfeito"	['gostei', 'bastante', 'tela', 'ótimo', 'tamanho', 'satisfatório', 'preço', 'celular', 'não', 'tenho', 'reclamar', 'satisfeito']
"celular bonito tudo certo não gostei não embalam telefone saco bolha caixa sem proteção telefone sem arranhão tudo certo"	['celular', 'bonito', 'tudo', 'certo', 'não', 'gostei', 'não', 'embalam', 'telefone', 'saco', 'bolha', 'caixa', 'sem', 'proteção', 'telefone', 'sem', 'arranhão', 'tudo', 'certo']
"dei presente esposa tudo certo"	['dei', 'presente', 'esposa', 'tudo', 'certo']
"achando muito bom fotos boas respostas rápidas biometria smart view bom carregador fraco demora carregar telefone"	['achando', 'muito', 'bom', 'fotos', 'boas', 'respostas', 'rápidas', 'biometria', 'smart', 'view', 'bom', 'carregador', 'fraco', 'demora', 'carregar', 'telefone']
"produto muito bom bateria dura bastante não gostei tamanho acostumado celular maior recomendo comprar sem medo"	['produto', 'muito', 'bom', 'bateria', 'dura', 'bastante', 'não', 'gostei', 'tamanho', 'acostumado', 'celular', 'maior', 'recomendo', 'comprar', 'sem', 'medo']
"produto design lindo pouca durabilidade bateria não devolvo necessidade aparelho"	['produto', 'design', 'lindo', 'pouca', 'durabilidade', 'bateria', 'não', 'devolvo', 'necessidade', 'aparelho']
"não tive sorte compra celular mês uso travando não obedece"	['não', 'tive', 'sorte', 'compra', 'celular', 'mês', 'uso', 'travando', 'não', 'obedece']
"qualidade câmeras deixa desejar básico dá gasto precisa adquirir separadamente"	['qualidade', 'câmeras', 'deixa', 'desejar', 'básico', 'dá', 'gasto', 'precisa', 'adquirir', 'separadamente']
"celular mediano esperava mais samsung travando app momento desigualdade tela reiniciar app"	['celular', 'mediano', 'esperava', 'mais', 'samsung', 'travando', 'app', 'momento', 'desigualdade', 'tela', 'reiniciar', 'app']
"produto após meses uso defeitos travamento acessos remotos apps ligações"	['produto', 'após', 'meses', 'uso', 'defeitos', 'travamento', 'acessos', 'remotos', 'apps', 'ligações']
"produto muito bom bateria dura bastante não gostei tamanho recomendo comprar sem medo"	['produto', 'muito', 'bom', 'bateria', 'dura', 'bastante', 'não', 'gostei', 'tamanho', 'recomendo', 'comprar', 'sem', 'medo']
"produto design lindo pouca durabilidade bateria não devolvo necessidade aparelho"	['produto', 'design', 'lindo', 'pouca', 'durabilidade', 'bateria', 'não', 'devolvo', 'necessidade', 'aparelho']
"configuração não adaptei a15 faz parte não ruim configuração sofrer já fui técnico samsung"	['configuração', 'não', 'adaptei', 'a15', 'faz', 'parte', 'não', 'ruim', 'configuração', 'sofrer', 'já', 'fui', 'técnico', 'samsung']
"produto errado negado reembolso descarada ilegal desprezo consumidor"	['produto', 'errado', 'negado', 'reembolso', 'descarada', 'ilegal', 'desprezo', 'consumidor']
"meu prestou 3 dias não carregou mais não aceitaram devolução fiquei prejuízo"	['meu', 'prestou', '3', 'dias', 'não', 'carregou', 'mais', 'não', 'aceitaram', 'devolução', 'fiquei', 'prejuízo']
"péssima qualidade 1 mês uso problema não carregando péssimo não compre"	['péssima', 'qualidade', '1', 'mês', 'uso', 'problema', 'não', 'carregando', 'péssimo', 'não', 'compre']
"celular bem lento devagar abrir apps executar tarefas simples copiar colar texto dois apps"	['celular', 'bem', 'lento', 'devagar', 'abrir', 'apps', 'executar', 'tarefas', 'simples', 'copiar', 'colar', 'texto', 'dois', 'apps']
"celular 1 mês uso apresentou problemas mandar assistência técnica"	['celular', '1', 'mês', 'uso', 'apresentou', 'problemas', 'mandar', 'assistência', 'técnica']

"celular bom áudio parece não proteção encaixaram cima metade proteção áudio cima"	['celular', 'bom', 'áudio', 'parece', 'não', 'proteção', 'encaixaram', 'cima', 'metade', 'proteção', 'áudio', 'cima']
"cinco dias uso defeito celular não primeira experiência desejar"	['cinco', 'dias', 'uso', 'defeito', 'celular', 'não', 'primeira', 'experiência', 'desejar']

APÊNDICE C – DATABASE DE TREINAMENTO

Avaliações Tokenizadas	Label
['qualidade', 'câmeras', 'deixa', 'desejar', 'básico', 'dá', 'gasto', 'precisa', 'adquirir', 'separadamente']	0
['dei', 'presente', 'esposa', 'tudo', 'certo']	1
['produto', 'design', 'lindo', 'pouca', 'durabilidade', 'bateria', 'não', 'devolvo', 'necessidade', 'aparelho']	0
['produto', 'errado', 'negado', 'reembolso', 'descarada', 'ilegal', 'desprezo', 'consumidor']	-1
['achando', 'muito', 'bom', 'fotos', 'boas', 'respostas', 'rápidas', 'biometria', 'smart', 'view', 'bom', 'carregador', 'fraco', 'demora', 'carregar', 'telefone']	1
['produto', 'muito', 'bom', 'bateria', 'dura', 'bastante', 'não', 'gostei', 'tamanho', 'acostumado', 'celular', 'maior', 'recomendo', 'comprar', 'sem', 'medo']	0
['celular', 'bom', 'áudio', 'parece', 'não', 'proteção', 'encaixaram', 'cima', 'metade', 'proteção', 'áudio', 'cima']	-1
['produto', 'design', 'lindo', 'pouca', 'durabilidade', 'bateria', 'não', 'devolvo', 'necessidade', 'aparelho']	1
['não', 'tive', 'sorte', 'compra', 'celular', 'mês', 'uso', 'travando', 'não', 'obedece']	0
['péssima', 'qualidade', '1', 'mês', 'uso', 'problema', 'não', 'carregando', 'péssimo', 'não', 'comprem']	-1
['produto', 'após', 'meses', 'uso', 'defeitos', 'travamento', 'acessos', 'remotos', 'apps', 'ligações']	0
['gostei', 'bastante', 'tela', 'ótimo', 'tamanho', 'satisfatório', 'preço', 'celular', 'não', 'tenho', 'reclamar', 'satisfeito']	1
['celular', 'bem', 'lento', 'devagar', 'abrir', 'apps', 'executar', 'tarefas', 'simples', 'copiar', 'colar', 'texto', 'dois', 'apps']	-1
['cinco', 'dias', 'uso', 'defeito', 'celular', 'não', 'primeira', 'experiência', 'desejar']	-1
['amplamente', 'elogiado', 'rapidez', 'ótima', 'câmera', 'bateria', 'durável', 'não', 'travar', 'durante', 'uso', 'apreciado', 'funções', 'nfc', 'smart', 'view', 'custo-benefício']	1
['meu', 'prestou', '3', 'dias', 'não', 'carregou', 'mais', 'não', 'aceitaram', 'devolução', 'fiquei', 'prejuízo']	-1

APÊNDICE D – DATABASE DE TESTE

Avaliações Tokenizadas	Label
['celular', '1', 'mês', 'uso', 'apresentou', 'problemas', 'mandar', 'assistência', 'técnica']	-1
['configuração', 'não', 'adapte', 'a15', 'faz', 'parte', 'não', 'ruim', 'configuração', 'sofrer', 'já', 'fui', 'técnico', 'samsung']	0
['celular', 'bonito', 'tudo', 'certo', 'não', 'gostei', 'não', 'embalaram', 'telefone', 'saco', 'bolha', 'caixa', 'sem', 'proteção', 'telefone', 'sem', 'arranhão', 'tudo', 'certo']	1
['produto', 'muito', 'bom', 'bateria', 'dura', 'bastante', 'não', 'gostei', 'tamanho', 'acostumado', 'celular', 'maior', 'recomendo', 'comprar', 'sem', 'medo']	1
['celular', 'mediano', 'esperava', 'mais', 'samsung', 'travando', 'app', 'momento', 'desigualdade', 'tela', 'reiniciar', 'app']	0