

MENSURANDO A SATISFAÇÃO DO USUÁRIO: UMA ABORDAGEM DE ANÁLISE DE SENTIMENTO PARA AVALIAÇÕES DE APPS FINANCEIROS

Fernando Henriques Neto, R.A.: 18.00931-0
Guilherme Sanches Rossi, R.A.: 19.02404-5
Matheus Coelho Rocha Pinto, R.A.: 20.00391-9
Pedro Henrique Sant Anna Hein, R.A.: 20.00134-7

RESUMO

Este estudo desenvolve um sistema automatizado para medir a satisfação dos usuários de aplicativos bancários, utilizando técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina para análise de sentimentos. Com dados coletados da Google Play Store, o sistema cria um ranking dos aplicativos mais bem avaliados em um intervalo de datas pré-definida pelo usuário, permitindo uma visão quantitativa e qualitativa sobre a experiência dos clientes. A interface em *Streamlit* e a execução em *Docker* facilitam a visualização e acessibilidade dos dados para gestores, que podem acompanhar a evolução semanal das percepções dos usuários. Este sistema pode orientar melhorias contínuas, ajudando instituições financeiras a adaptarem-se melhor às demandas do mercado digital.

Palavras-chave: Análise de Sentimentos; Aplicativos Bancários; Processamento de Linguagem Natural (PLN); Satisfação do Usuário.

ABSTRACT

This study develops an automated system to measure user satisfaction with banking applications, using Natural Language Processing (NLP) techniques and machine learning for sentiment analysis. With data collected from the Google Play Store, the system creates a weekly ranking of the highest-rated apps, providing both quantitative and qualitative insights into customer experience. The interface built in *Streamlit* and deployed with *Docker* facilitates data visualization and accessibility for managers, who can monitor weekly changes in user perceptions. This system provides strategic insights that guide continuous improvements, helping financial institutions better adapt to the demands of the digital market.

Keywords: Sentiment Analysis; Banking Applications; Natural Language Processing (NLP); User Satisfaction.

1 INTRODUÇÃO

O avanço da tecnologia tem transformado a relação dos consumidores com instituições bancárias, promovendo conveniência e acessibilidade através de aplicativos e serviços de mobile banking. Para aqueles que já enfrentaram a

necessidade de resolver questões bancárias em agências físicas, a chegada dos aplicativos representa uma mudança significativa, oferecendo rapidez e flexibilidade diretamente nos dispositivos móveis. Apesar das facilidades que o mobile banking traz, não são todos que são desenvolvidos com foco na experiência de usuário, o que pode gerar uma barreira de uso (VITORINO, 2013).

Esse cenário de digitalização é confirmado por uma pesquisa recente do Instituto Ipsos, encomendada pelo Nubank, que revela que 38% dos brasileiros preferem os canais digitais às agências bancárias, com uma tendência ainda mais forte entre a Geração Z, dos quais 16% nunca usaram o caixa de uma agência (CAMPOS, 2024). Embora a adoção digital tenha avançado em todas as faixas etárias, diferenças de comportamento entre gerações evidenciam desafios e oportunidades na adaptação das experiências bancárias.

A avaliação das opiniões dos usuários sobre aplicativos é crucial para que os desenvolvedores entendam expectativas, identifiquem falhas e direcionem melhorias. No entanto, devido ao alto volume de avaliações, o processo manual é impraticável. Nesse sentido, sistemas automatizados de análise de sentimentos, como o desenvolvido por Pereira (2015), surgem como uma solução para processar feedback em larga escala e destacar insights valiosos para a tomada de decisão.

Este estudo visa explorar a "porcentagem de positividade" nas avaliações de aplicativos bancários na Google Play Store, identificando a proporção de comentários positivos em um intervalo de datas pré-definido pelo usuário. Utilizando técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina, o objetivo é criar um ranking dos aplicativos mais bem avaliados, fomentando discussões sobre como aprimorar a experiência do usuário no setor financeiro. Dessa forma, espera-se contribuir para um entendimento mais profundo das expectativas dos usuários e apoiar as instituições financeiras na adaptação aos feedbacks dos clientes.

2 EXTRAÇÃO DOS DADOS

Antes de iniciar a construção do sistema de avaliação de satisfação de usuário, é necessário realizar a extração dos dados. Para isto, utilizamos a biblioteca *google_play_scraper*, que facilita o acesso aos comentários dos usuários diretamente da plataforma. A primeira etapa consistiu na definição dos aplicativos a serem analisados, representados pelos seus respectivos IDs na Play Store. Foram selecionados os principais aplicativos de bancos brasileiros, sendo eles: Itaú, Nubank, Banco do Brasil, Bradesco, Santander, C6 bank, XP investimentos e Banco Inter.

A extração foi realizada aplicando um filtro de idioma e país, garantindo que apenas avaliações em português e feitas por usuários brasileiros fossem coletadas. Optamos por classificar as avaliações pela ordem mais recente para obter uma visão atualizada da percepção dos usuários sobre os aplicativos e conseguir filtra-los pelas datas. A coleta foi limitada a um número máximo de 8.000 avaliações por aplicativo, totalizando 64.000 avaliações, o que nos proporcionou uma amostra ampla e variada para a análise de sentimentos.

Esta extração nos gerou um arquivo (*app_review.csv*), e as principais colunas que utilizaremos para este estudo são, o ID de cada banco, os comentários, a nota real que o usuário selecionou na Play Store e a data que o comentário foi feito. Essas informações nos permitirá uma análise mais profunda dos sentimentos conforme serão detalhados nas etapas seguintes deste projeto.

3 DESENVOLVIMENTO

Para a construção do sistema de **ranking dos aplicativos bancários mais bem avaliados** na Google Play Store, utilizamos *python* com a biblioteca *streamlit*, que permite criar uma interface interativa para visualizar e classificar os aplicativos com base na avaliação de seus usuários, em seguida realizamos o *upload* em um ambiente Docker. As principais etapas do desenvolvimento foram: importação e preparação dos dados, classificação das avaliações, cálculo de porcentagens de sentimentos, criação do ranking, implementação da interface visual e contenção com *Docker*.

Processamento de Dados e Aplicação do LLM BERT

Após a extração dos dados aplicamos o modelo *FinBERT-PT-BR*, um modelo especializado em análise de sentimentos financeiros e ajustado para o português, que nos permitiu classificar os comentários em três categorias: **Positivo**, **Negativo** e **Neutro**.

Para isso, utilizamos a biblioteca *Transformers* e carregamos o *tokenizer* e o modelo pré-treinado do *FinBERT*. Cada comentário foi processado e os resultados foram incluídos e armazenados em um novo CSV (*df_short.csv*), facilitando seu uso nas etapas seguintes.

Modelos de Classificação com Machine Learning

Após a classificação inicial com *LLM BERT*, implementamos diversos modelos de Machine Learning para prever os sentimentos dos comentários com base nas variáveis processadas. Os modelos selecionados são:

- *Regressão Logística*,
- *K-Nearest Neighbors (KNN)*,
- *Random Forest*,
- *BernoulliNB*, *GaussianNB* e *MultinomialNB* para *Naive Bayes*.

Primeiramente, utilizamos o *TfidfVectorizer* para transformar os textos em representações numéricas, capturando as principais palavras presentes nos comentários.

Os dados foram então divididos em conjuntos de treinamento e teste para as variáveis dependentes *sentiment* (*avaliações reais dos usuários*) e *label* (*avaliações classificadas pelo LLM Bert*), possibilitando a avaliação dos modelos para as duas colunas. Cada modelo foi treinado e avaliado com a métrica de **acurácia**. Entre os modelos avaliados, destacamos aquele que obteve a melhor acurácia para ambas as variáveis dependentes, o *RandomForest* (0.86 *sentiment* e 0.85 *label*).

Com isso, geramos um novo arquivo CSV com as seguintes colunas, o ID de cada banco, os comentários dos usuários, as avaliações reais dos usuários, as avaliações previstas do LLM Bert, a predição do *RandomForest* em cima das avaliações reais e por fim a predição do *RandomForest* das avaliações do LLM Bert.

Visualização e Análise de Resultados

Para a análise dos resultados, desenvolvemos uma interface interativa com o *Streamlit*. Esta interface permite ao usuário selecionar o período de interesse para análise. Através de funções desenvolvidas para calcular a porcentagem de sentimentos positivos, foi possível criar um *ranking* dos bancos mais bem avaliados, facilitando a comparação entre o sentimento real e as previsões dos modelos.

As principais funcionalidades da interface incluem:

- **Seleção de Período:** através de um calendário interativo, os usuários podem definir o intervalo de datas para a análise das resenhas.
- **Ranking dos Bancos:** os bancos foram ordenados pela porcentagem de sentimentos positivos do LLM Bert. As diferenças entre as previsões, os valores reais e a acurácia de cada modelo são exibidas, permitindo uma análise qualitativa da precisão dos modelos.

O projeto está disponível no seguinte repositório do github:

https://github.com/PedroHein/T6_CD2

4 CONSIDERAÇÕES FINAIS

Este estudo propôs um sistema automatizado de análise de sentimentos para avaliar a satisfação dos usuários de aplicativos bancários, aplicando técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina. A coleta e análise das avaliações na Google Play Store permitiu identificar os principais pontos de satisfação e insatisfação em relação a cada aplicativo, criando um ranking por período selecionado pelo usuário dos bancos mais bem avaliados. Essa abordagem possibilitou que grandes volumes de feedback fossem processados e interpretados automaticamente, fornecendo uma visão abrangente da experiência dos usuários em tempo real.

A interface visual, desenvolvida com Streamlit e executada em Docker, tornou os dados acessíveis e organizados, facilitando a consulta de gestores. Essa estrutura fornece informações estratégicas, como a evolução por período de tempo da percepção dos usuários e o desempenho relativo dos principais bancos, ajudando as instituições financeiras a direcionarem melhorias. A análise com percentuais de positividade oferece uma métrica quantitativa que complementa o feedback qualitativo, permitindo uma interpretação prática das expectativas e insatisfações.

No entanto, uma das conclusões que chegamos é que a acurácia (49,7%) do modelo de análise de sentimentos (LLM Bert) ainda enfrenta desafios, pois há variáveis que ele não consegue interpretar adequadamente. Por exemplo, comentários irônicos e contradições (como comentários positivos acompanhados de avaliações baixas e vice-versa). Esses fatores diminuem a precisão da análise e indicam a necessidade de aprimoramentos futuros para lidar com nuances mais complexas na comunicação dos usuários.

Por fim, devemos ressaltar a acurácia que identificamos entre o modelo de classificação (RandomForest) e as avaliações, tanto reais, quanto a do LLM Bert. Percebemos que a acurácia entre o modelo e as avaliações reais ficou muito mais baixa, quando comparada com o modelo e as avaliações do LLM Bert. Isso se deve ao fato de que o modelo de classificação RandomForest está aprendendo a partir dos padrões identificados nas avaliações do LLM Bert, que apresentam uma certa consistência nos rótulos e podem conter menos variações em relação às avaliações reais dos usuários.

5 REFERÊNCIAS BIBLIOGRÁFICAS

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2017). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems*, 28(2), 15-21.

CAMPOS, Álvaro. Quase 40% dos entrevistados não foram a uma agência bancária nos últimos seis meses, mostra pesquisa. VALOR, São Paulo, 22/10/2024. Disponível em: <https://valor.globo.com/financas/noticia/2024/10/22/quase-40percent-dos-entrevistados-nao-foram-a-uma-agencia-bancaria-nos-ultimos-seis-meses-mostra-pesquisa.ghml>. Acesso em: 03/11/2024.

OLIVEIRA, Rogério de. CASE: Avaliação de aplicativos de comida do Google Play. Disponível em: https://colab.research.google.com/github/Rogério-mack/IMT_Ciencia_de_Dados/blob/main/IMT_Sentiment_Analysis_TFIDF_solucao.ipynb#scrollTo=WtiXMksjxBK4. Acesso em: 07/10/2024.

PEREIRA, Renan Andrade. DESENVOLVIMENTO DE UM SISTEMA DE ANÁLISE DE SENTIMENTO PARA TEXTOS DE AVALIAÇÃO DE APLICATIVOS. 2015.

VITORINO, Carla. Análise da usabilidade do App Mobile Banking do Banco Santander. **PUC-RIO. Relatório (Especialização)**, 2013.