

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS

PAULO AUGUSTO FERNANDES DE SOUZA

Mineração de emoção para avaliar a experiência de  
usuário em serviços bancários

São Carlos

2022

PAULO AUGUSTO FERNANDES DE SOUZA

Mineração de emoção para avaliar a experiência de  
usuário em serviços bancários

Monografia apresentada ao Curso de Engenharia de Produção, da Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Engenheiro de Produção.

Orientadora: Prof. Dr. Carina Campese

Coorientadora: Me. Adriana Hofmann  
Trevisan

São Carlos

2022



AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTA TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

S719m Souza, Paulo  
Mineração de emoção para avaliar a experiência de usuário em serviços bancários / Paulo Souza; orientadora Carina Campese; coorientadora Adriana Trevisan. São Carlos, 2022.

Monografia (Graduação em Engenharia de Produção) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2022.

1. Mineração de emoção. 2. Experiência do usuário. 3. Análise de mídias sociais. I. Título.

## FOLHA DE APROVAÇÃO

<b>Candidato:</b> Paulo Augusto Fernandes de Souza
<b>Título do TCC:</b> Mineração de emoção para avaliar a experiência de usuário em serviços bancários
<b>Data de defesa:</b> 28/01/2022

Comissão Julgadora	Resultado
Professora Doutora Carina Campese (orientadora)	APROVADO
Instituição: EESC - SEP	
Professora Associada Daisy Aparecida do Nascimento Rebelatto	APROVADO
Instituição: EESC - Departamento de Engenharia de Produção	
Pesquisadora Adriana Hofmann Trevisan	APROVADO
Instituição: EESC - Departamento de Engenharia de Produção	

Presidente da Banca: **Professora Doutora Carina Campese**

## RESUMO

SOUZA, P. A. F. **Mineração de emoção para avaliar a experiência de usuário em serviços bancários**. 2022. Monografia (Trabalho de Conclusão de Curso) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2022.

O segmento bancário tem uma grande relevância na economia brasileira e tem sofrido grandes mudanças na última década. O surgimento de empresas financeiras totalmente digitalizadas, as chamadas *fintechs*, gerou um novo ramo no segmento bancário e uma nova maneira de consumir produtos financeiros. Esta transformação no mercado gerou também uma experiência diferente para os usuários das plataformas das *fintechs*, e com isso a discussão se esta experiência seria mais satisfatória do que a tradicional. Tendo em vista que a experiência do cliente pode ir desde a usabilidade técnica de um produto até a percepção particular que o usuário tem de uma empresa, surge em conjunto a necessidade de se avaliar esta experiência a partir de perspectivas mais variadas. Portanto, esta pesquisa tem o objetivo de avaliar a experiência do usuário de serviços bancários, sob a ótica da comparação entre *fintechs* e bancos tradicionais, por meio de dados extraídos da rede social Twitter. Esta pesquisa comparou a maior instituição privada de cada ramo, *fintech* e banco tradicional, do Brasil. Estes dados foram interpretados por meio de algoritmos de mineração de textos e mostraram que os usuários da *fintech* se expressam mais positivamente sobre a empresa do que os usuários do banco tradicional. Junto a esta análise, foi realizada uma análise exploratória em outras vias de avaliação de experiência consolidadas que indicaram também uma vantagem de experiência para as *fintechs* guiadas principalmente pelo engajamento social que os clientes têm no relacionamento com as empresas digitais.

Palavras-chave: Mineração de emoção. Experiência do usuário. Análise de mídias sociais.



## ABSTRACT

SOUZA, P. A. F. **Emotion mining to evaluate the user experience in bank services.** 2022. Monografia (Trabalho de Conclusão de Curso) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2022.

The banking segment has great relevance in the Brazilian economy and has undergone major changes in the last decade. The emergence of fully digitized financial companies, the so-called *fintechs*, has generated a new branch in the Brazilian banking segment and a new way of consuming financial products. This transformation in the market also generated a different experience for the users of *fintech* platforms, and with that the discussion whether this experience would be more satisfactory than the traditional one. Considering that the customer experience can range from the technical usability of a product to the particular perception that the user has of a company, the need to evaluate this experience from more varied perspectives arises together. Therefore, this research aims to evaluate the user experience of banking services, from the perspective of the comparison between *fintechs* and traditional banks, through data extracted from the social network Twitter. This research compared the largest private institution in each branch, *fintech* and traditional bank, in Brazil. These data were interpreted using text mining algorithms and showed that *fintech* users express themselves more positively about the company than traditional bank users do. Along with this analysis, an exploratory analysis was carried out on other consolidated ways of evaluating experience that also indicated an experience advantage for *fintechs* guided mainly by the social engagement that customers have in the relationship with digital companies.

Key-words: Emotion mining. User experience. Social media analytics.



# Sumário

RESUMO .....	5
ABSTRACT .....	7
Lista de Figuras.....	10
1. INTRODUÇÃO .....	11
1.1. Contexto e Justificativa da pesquisa .....	11
1.2. Objetivos da pesquisa .....	13
2. REVISÃO BIBLIOGRÁFICA.....	13
2.1. A experiência do usuário e o relacionamento com empresas.....	13
2.2. Redes sociais na avaliação da experiência do usuário.....	15
2.3. Twitter como fonte de dados.....	16
2.4. Mineração e análise textual em redes sociais .....	17
3. METODOLOGIA .....	18
3.1. Etapa I – Revisão da literatura .....	19
3.2. Etapa II – Coleta e processamento dos dados .....	20
3.3. Etapa III – Análise dos resultados .....	24
4. Resultados.....	26
4.1. Análise de Sentimentos dos dados coletados do Twitter .....	27
4.2. Análise Exploratória .....	32
4.3. Comparação entre as vias .....	40
5. Conclusão .....	41
6. Referências bibliográficas .....	42



## Lista de Figuras

Figura 1 - Etapas da Metodologia (Fonte: Autor, 2021).....	19
Figura 2 - Código de acesso (Fonte: Autor, 2021).....	20
Figura 3 - Código de busca (Fonte: Autor, 2021) .....	21
Figura 4 - Código de preparação de textos (Fonte: Autor, 2021).....	23
Figura 5 - Código de construção do classificador (Fonte: Autor, 2021) .....	24
Figura 6 - Código de atribuição de notas aos textos (Fonte: Autor, 2021) .....	24
Figura 7 - Código de construção da nuvem de palavras (Fonte: Autor, 2021).....	25
Figura 8 - Gráfico de frequência dos sentimentos nos textos (Fonte: Autor, 2022) .....	27
Figura 9 - Nuvem de palavras sobre a <i>fintech</i> Nubank (Fonte: Autor, 2022) .....	29
Figura 10 - Nuvem de palavras sobre o banco Bradesco (Fonte: Autor, 2022) .....	29
Figura 11 - Avaliações do banco no Google Play (Fonte: Google Play, 2021) .....	33
Figura 12 - Avaliações do banco Bradesco na loja da APPs da Apple (Fonte: Loja de Apps Apple, 2021) .....	33
Figura 13 - Figura 13. Avaliações da <i>fintech</i> no Google Play Apple (Fonte: Google Play, 2021) .....	34
Figura 14 - Avaliações da <i>fintech</i> na loja de APPs da Apple (Fonte: Loja de Apps Apple, 2021) .....	34
Figura 15 - Buscas no Google Trends (Fonte: Google Trends, 2021) .....	35
Figura 16 – Avaliações de clientes no Reclame Aqui (Fonte: Reclame Aqui, 2021).....	37
Figura 17 - Ranking de interações dos clientes com empresas financeiras pela SoluCX (Fonte: SoluCX, 2021) .....	38
Figura 18 - Ranking de NPS das empresas financeiras (Fonte: SoluCX, 2021) .....	38
Figura 19 - Número de clientes do banco Bradesco no Instagram e Twitter (Fonte: Instagram, 2021; Twitter, 2021) .....	39
Figura 20 - Número de clientes da <i>fintech</i> Nubank no Instagram e no Twitter (Fonte: Instagram, 2021; Twitter, 2021) .....	39

## 1. INTRODUÇÃO

### 1.1. Contexto e Justificativa da pesquisa

No mundo inteiro, o setor bancário movimenta trilhões de dólares e em qualquer país este segmento tem importância vital. Neste setor, há empresas gigantes que tiveram grandes lucros ao longo do tempo e, com a sua rentabilidade, sempre houve o interesse por meio de outras empresas de competir por este mercado de tamanha atração (CRUZ-GARCÍA; FERNÁNDEZ DE GUEVARA; MAUDOS, 2021). No Brasil, dados do Banco Central (BACEN, 2020) revelam que o setor bancário nacional ainda é extremamente concentrado, uma vez que o grupo formado pelos cinco maiores bancos possui no agregado mais de 80% dos ativos totais do segmento bancário comercial. Porém, houve de 2018 para 2019 um crescimento de 29% na quantidade de empresas financeiras digitais, e um aumento de mais de 42% dos ativos controlados por instituições de pagamentos de 2017 para 2019 (BACEN, 2020).

Nos últimos anos, surgiram empresas com alta tecnologia para entrar com um modelo diferente neste mercado. Estas empresas com maior foco em tecnologia que competem no mercado bancário são nomeadas *fintechs*. As *fintechs* podem ser definidas como um neologismo que descreve a conexão do moderno, como relacionado à internet, tecnologias (por exemplo, computação em nuvem, internet móvel) e atividades de negócios típico da indústria de serviços financeiros (GOMBER; KOCH; SIERING, 2017). A visão intrínseca das *fintechs* é a inserção de maneiras mais baratas de superar os atritos de contratação financeira, por excluir a interação presencial e estrutura física, e reduzir custos dos serviços financeiros para melhorar a experiência do consumidor por meio de uma interface completamente digital (THAKOR, 2020).

Inicialmente, as empresas de serviços financeiros digitais emergiram como startups para oferecer serviços financeiros melhores (MILIAN; SPINOLA; CARVALHO, 2019). A grande agilidade com informações, a personalização e efetividade do serviço oferecido por essas empresas, forçou os players tradicionais a adotar as tecnologias para manterem a competitividade (AZARENKOVA et al., 2018). Por isso, de acordo com Vardomatskya, Kuznetsova, Plotnikov (2021), o segmento bancário, como os demais, experimentou os efeitos da “quarta revolução tecnológica”<sup>1</sup>. Neste sentido, há uma grande diferença na maneira em que os bancos tradicionais se relacionam com os seus clientes em comparação com o modo que as *fintechs* o fazem, já que estas baseiam suas operações em plataformas digitais (VARDOMATSKYA; KUZNETSOVA; PLOTNIKOV, 2021). Os bancos tradicionais também utilizam os meios digitais, como os aplicativos mobile, mas o fazem apenas como suporte às operações dos clientes, como transações financeiras ou o uso de serviços bancários enquanto mantêm interações físicas (OWUSU KWATENG; OSEI ATIEMO; APPIAH, 2019).

Desta forma, as duas categorias de empresas possuem processos diferentes de como se relacionar com o cliente pelos meios de comunicação digitais (MILIAN; SPINOLA; CARVALHO, 2019). Segundo Kapoor, Vij (2019) cada mudança na plataforma de interação pode gerar uma percepção diferente por parte do usuário. Os autores ainda ressaltam que a avaliação dos clientes sobre as plataformas influencia positivamente a recomendação dos canais distribuidores para novos clientes. Logo, uma boa percepção dos apps e mídias digitais alavancam o alcance da marca para o mercado como um todo (KAPOOR; VIJ, 2020). Da mesma forma, Garzaro, Varotto, Pedro (2020) afirmam que a qualidade que estes meios

---

<sup>1</sup> “Indústria 4.0 é uma filosofia de manufatura que inclui sistemas de automação modernos com uma autonomia de nível intenso, trocas de dados flexíveis e eficazes que permitem a implementação de tecnologias de produção de última geração, inovação em design e mais pessoal e mais ágil na produção, bem como produtos personalizados.” (OZTEMEL; GURSEV, 2020, p. 166)

possuem com relação à experiência do usuário influencia diretamente em seu engajamento e lealdade à empresa que fornece este serviço. Desta maneira, aplicativos e plataformas que oferecem uma experiência mais sofisticada e com maior presença social por parte da interface alcançam maior engajamento dos clientes com a marca, o que influencia diretamente o retorno que o cliente traz, e a lealdade dos consumidores com relação às demais ofertas que ele possui no mercado (GARZARO; VAROTTO; PEDRO, 2020).

Para Lee et al. (2021), os meios de interações virtuais deram às pessoas uma plataforma para dar feedback e ampliou a capacidade do cidadão de demonstrar sua satisfação. Ainda segundo os autores, este processo acontece sem qualquer burocracia ou preparação de pesquisa por meio de alguma instituição, logo fornece uma perspectiva que os demais canais não fazem (LEE et al., 2021). Atualmente, é comum que as pessoas se comuniquem e até se expressem abertamente sobre algum tema determinado nas redes sociais, como Twitter, Facebook e Instagram (SADAGHEYANI; TATARI, 2020). Desta forma, muitos dados sobre as opiniões dos clientes podem ser coletados nestes meios de comunicação. Estas plataformas podem revelar a qualidade dos serviços prestados, e além disso as redes sociais podem ter um papel muito importante sobre a performance destas empresas uma vez que os consumidores conseguem exercer influência nos demais clientes por meio de suas opiniões (JALILVAND et al., 2017).

Com as diversas tecnologias que emergiram no cenário recente, grupos distintos de pessoas passaram a usar meios de comunicação diferentes para se expressar, interagir e relatar problemas; alguns grupos, concentram sua comunicação nas redes sociais (KARAM; HUSSEIN; GHARIB, 2021). Neste contexto, muitos usuários nem chegam a interagir com outros meios. Diante disso, surge a necessidade de avaliar a experiência do usuário através da utilização de dados coletados das redes sociais, já que as empresas podem otimizar seus negócios e melhorar a experiência dos clientes pelo conhecimento de suas opiniões (FERNANDES et al., 2021).

Diversas instituições mantêm apenas a avaliação da satisfação dos seus clientes por meio das vias formais como ouvidoria interna, sites de avaliação credenciados (consumidor.gov, Reclame Aqui, BACEN) ou pesquisas consolidadas de satisfação (NPS). Estas vias de reclamações possuem um imenso valor de conhecimento, pois estão padronizadas e adaptadas para receber opiniões de usuários. Mas elas coletam apenas as visões de clientes que conhecem estes canais e que se expressam sobre experiência pontuais negativas. No caso das pesquisas consolidadas, os clientes dão uma nota, mas poucas vezes dão informações relevantes para que se entenda o motivo da avaliação. Logo, as entidades que consultam estas vias perdem os detalhes de uma variedade maior de clientes, bem como de expressões mais completas e mais naturais sobre a experiência com a empresa. Nas redes sociais, os usuários vão se expressar de maneira mais livre e com mais abertura pois a sua avaliação será compartilhada com as pessoas de um vínculo mais próximo (ZHAN et al., 2021).

Portanto, para coletar estas informações, é necessário o uso de ferramentas que nem sempre são difundidas no ambiente acadêmico ou empresarial. Desta forma, pelo desconhecimento destes instrumentos, muitas instituições ignoram ou não compreendem as opiniões dos grupos que migraram sua comunicação para as mídias sociais (JOHANN; WOLF; GODULLA, 2021). Logo, esta pesquisa almeja avaliar a diferença na experiência do usuário de acordo com o modelo de instituição bancária por meio de dados coletados nas redes sociais<sup>2</sup>. Esta

---

<sup>2</sup> Não será avaliado nesta pesquisa a representatividade dos dados das redes sociais diante da população de clientes dos serviços bancários; será, pois, avaliado apenas a experiência dos que se expressam. Em outras palavras, não será feita uma verificação se o volume de usuários que se manifestam sobre os bancos nas mídias são uma amostra significativa dentro do grupo de clientes dos bancos.

análise será feita por meio de algoritmos computacionais de interpretação e análise de textos aliados a resultados que outras vias já realizam.

## **1.2. Objetivos da pesquisa**

O objetivo geral deste estudo é avaliar a experiência do usuário com relação ao uso do aplicativo mobile de serviços bancários. A avaliação será baseada em mineração textual de emoções de conteúdo gerado pelas redes sociais. Esse objetivo se desdobra em dois objetivos específicos:

- O primeiro objetivo é comparar a qualidade da experiência do usuário de um banco tradicional com um banco digital por meio de algoritmos de interpretação de textos. Desta maneira, analisar de maneira quantitativa e qualitativa o que os usuários expressam de positivo e negativo nas redes sociais a respeito de cada categoria.
- O segundo objetivo é comparar estes resultados com outras avaliações consolidadas para contextualizar a análise e também revelar o valor adicional que ela pode trazer.

## **2. REVISÃO BIBLIOGRÁFICA**

A revisão da literatura realizada consistiu de quatro partes que deram base à pesquisa. Estas partes buscam construir a relação entre os diversos campos abordados, bem como trazer um contexto profundo daquilo que é estudado em cada área. Assim, a revisão da literatura se inicia com o estudo da experiência do usuário e o relacionamento destas pessoas com as empresas. Nesta primeira parte, estudou-se quais são os fatores que podem influenciar na experiência do cliente e como isso afeta a relação dele com a empresa. A partir destas informações, pudemos entender qual a maneira de construir uma boa experiência para o cliente. Na segunda parte, buscou-se entender o quanto as redes sociais podem ser efetivas em revelar a qualidade da experiência do usuário a partir daquilo que ele publica. Assim, foi realizado um estudo sobre o alcance destas plataformas e também como seria feito esta avaliação para averiguar inclusive o relacionamento dos clientes com empresas.

Em seguida, as próximas duas partes examinaram a parte técnica de como avaliar esta experiência do usuário a partir dos dados coletados nas redes sociais. Desta forma, a terceira parte buscou referências sobre a plataforma específica usada nesta pesquisa, o Twitter, e a possibilidade dos dados fornecidos por ela serem eficazes. Durante esta busca, foram usados diversos estudos como referência do uso desta mídia social. Em sequência, estudou-se a maneira de analisar os dados para obter informações concretas a respeito do objeto de estudo. Assim, a quarta parte trouxe estruturas montadas por pesquisadores de como interpretar textos de maneira escalável com o uso de algoritmos computacionais como o que foi utilizado nesta pesquisa. Assim,

### **2.1. A experiência do usuário e o relacionamento com empresas**

A experiência do usuário (user experience, UX) é um campo de estudo dentro de ergonomia que busca descrever as interações de uma pessoa com alguma interface (BRADLEY et al., 2021). A ISO (2019) define experiência do usuário como “percepções e respostas da pessoa resultantes do uso e / ou uso antecipado de um produto, sistema ou serviço”. Porém, além da definição objetiva, também podemos entender que a experiência vai além do uso físico de algum objeto e tampouco é restrita ao momento do uso (HASSENZAHN; TRACTINSKY, 2006). Segundo Vemeeren et al (2010), é extremamente importante entender como a experiência se desenvolve com o tempo, uma vez que as interações repetidas com um produto ou serviço podem mudar esta percepção que o usuário possui. Os autores ainda afirmam que a UX é mais abrangente do que apenas a usabilidade de um produto. A

usabilidade seria a facilidade de completar uma tarefa enquanto a UX também contempla a percepção e as sensações que se originam da interação como um todo (VERMEEREN et al., 2010).

De acordo com a pesquisa de Law et al (2009), existe uma grande complexidade em definir o termo UX, pois há uma grande diversidade de fatores que podem afetar a experiência do cliente. Desta maneira, muitos especialistas concordam que a UX se relaciona profundamente até com a imagem ou percepção que o cliente possui da empresa e não apenas do usuário (LAW et al., 2009). Desta forma, a experiência não se forma apenas da interação, mas também de valores ou sentimentos únicos que cada usuário traz, o que faz com que cada experiência seja distinta mesmo diante de um produto idêntico (HASSENZAHL; TRACTINSKY, 2006).

As diferenças dentro da própria base de usuários devem ser consideradas como fatores relevantes para se avaliar a experiência de um produto ou serviço (CHAWLA; JOSHI, 2018). Desta maneira, nenhuma empresa deve buscar atender apenas um segmento de clientes, ainda que seja o principal. É, pois, vital que as empresas consigam aferir e atender a amplitude de clientes que possuem, uma vez que esta diversidade gerará perspectivas distintas e, portanto, percepções diferentes da experiência (BRADLEY et al., 2021). Logo, a construção da plataforma ou produto que será exposto ao cliente deve buscar uma flexibilidade para atender requisitos variados de diferentes categorias de usuários (BRADLEY et al., 2021; CHAWLA; JOSHI, 2018).

Para esta construção de produtos e serviços, é extremamente relevante tomar como base o Design Centrado no Usuário (UCD) ou o Design Centrado no Humano (HCD). O termo UCD descreve uma filosofia de projetar sistemas para o uso das pessoas, com o possível envolvimento delas no planejamento, para atingir um estágio de satisfação e qualidade adequado (ABRAS; MALONEY-KRICHMAR; PREECE, 2004). Porém, Campese, Amaral e Mascarenhas (2020) discutem que a literatura ainda possui dificuldade em estipular uma definição adequada para os termos UCD/HCD. Devido à maior difusão na academia e também em buscas na internet, os autores recomendam a preferência no uso do termo HCD com a seguinte definição:

“O HCD é uma abordagem de envolvimento do usuário que compreende princípios que orientam o estabelecimento de objetivos e a realização de atividades, permitindo que a informação e o conhecimento do usuário sejam obtidos e incorporados no processo de design, utilizando métodos e técnicas identificadas no corpo de conhecimentos desta área” (CAMPESE; AMARAL; MASCARENHAS, 2020, p. 47).

Portanto, a definição vai além de apenas atender as necessidades do utilizador, mas também ressalta a importância de incorporar a voz do usuário no processo de desenvolvimento. Assim, torna-se essencial coletar feedback dos usuários, bem como envolvê-los no processo de design com uma maneira estruturada de coleta de informações e avaliações (KUJALA, 2008).

Todo este planejamento visa aprimorar a qualidade da experiência. Para isso, incrementos na usabilidade do cliente por meio de design, informação e eficiência podem fazer com que os usuários tenham uma experiência melhor, baseado na avaliação dos clientes (KAPOOR; VIJ, 2020). Isto melhora a atratividade para o cliente e, assim, fomenta o maior uso do produto pelo usuário (MUSLIM et al., 2019). Um fluxo de atividades bem construído pode aumentar significativamente a percepção do usuário e estimular um uso mais intenso de uma determinada solução (DOHERTY; SORENSON, 2015). Porém, o aumento na atratividade não se dá apenas por priorizar uma boa usabilidade, mas uma experiência com excelência aumenta a confiança e lealdade do cliente no relacionamento com a marca da empresa (JAMSHIDI et al., 2018). Segundo Garzaro, Varotto e Pedro (2020), a construção de uma boa

relação do cliente com a marca aumenta a confiança e o engajamento tanto com o produto ou serviço específico que iniciou o relacionamento, como também com toda a gama que a marca consegue oferecer pela lealdade que conquistou. Torna-se, assim, imprescindível que a companhia construa sua presença também nas redes sociais para alavancar a sua relação com os clientes e aprimorar a experiência oferecida (GARZARO; VAROTTO; PEDRO, 2020).

## **2.2. Redes sociais na avaliação da experiência do usuário**

As redes sociais são serviços de interação pela internet formadas por bases de dados. Por meio de um conjunto de funcionalidades virtuais que mediam as relações dos usuários, estas redes dão a possibilidade para que eles publiquem informações pessoais e indiquem preferências (SADAGHEYANI; TATARI, 2020). Estes meios de comunicação já estão extremamente integrados na rotina das pessoas; estudos da Organização para a Cooperação e Desenvolvimento Econômico (OCDE, 2017) apontam que 95% dos adolescentes de países que integram a organização interagem com redes sociais. O estudo também aponta que, na média, estudantes usam as redes mais de duas horas por dia (OECD, 2017).

As redes sociais têm se tornado extremamente difundidas e relevantes na vida das pessoas. Segundo Karam, Hussein, Gharib (2021), muitas pessoas usam estes meios para relatar situações problemáticas ou crises que atingem a sua comunidade. Deste modo, os autores ainda afirmam que as redes sociais se tornaram não apenas um meio de interação, mas um grande reservatório de dados das pessoas, bem como uma maneira de os usuários informarem as suas visões sobre determinado objeto (KARAM; HUSSEIN; GHARIB, 2021). Com a disposição dos usuários e a vasta acessibilidade às plataformas de comunicação, várias oportunidades emergem para a aplicação destes dados em diversas áreas (ZHAN et al., 2021). Da mesma forma, diversas pesquisas mostram a acessibilidade destes dados para serem usados de maneira analítica, e também revelam metodologias para que estas análises obtenham resultados relevantes (TENG; KHONG, 2021).

As avaliações fornecidas pelos clientes por meio das redes sociais e plataformas digitais podem ter uma precisão significativa para desvendar e implementar melhorias em serviços caso sejam estruturadas de maneira adequada (FERNANDES et al., 2021). Logo, a busca por métodos que extraiam de modo adequado estas informações têm crescido exponencialmente (KARAM; HUSSEIN; GHARIB, 2021). Segundo Lee et al. (2021), inúmeras instituições públicas e privadas já obtiveram métodos validados para coletar e avaliar de maneira proveitosa uma percepção detalhada e completa das pessoas sobre determinadas práticas de instituições.

O vasto acervo disponível online tem grande utilidade para avaliar serviços ou produtos e explorar melhorias (LEE et al., 2021). Segundo Johann, Wolf, Godulla (2021), os usuários das redes sociais vão além de apenas externalizar pensamentos, mas também utilizam a plataforma para construir e solidificar relacionamentos. Os autores ainda afirmam que este relacionamento não se restringe apenas às pessoas, mas a intensidade da conexão virtual afeta diretamente a construção de relações com empresas e instituições (JOHANN; WOLF; GODULLA, 2021). Assim, os usuários destas mídias são intensamente influenciados pelas opiniões e expressões dos seus pares sobre um determinado objeto (JALILVAND et al., 2017).

Existe uma grande variedade de redes sociais, dentre elas Facebook, Instagram, MySpace, Twitter e YouTube; e grande parte dos jovens mantém um uso tão intenso delas que muitas vezes pode ser prejudicial (MCNAMEE; MENDOLIA; YEROKHIN, 2021). Cada uma delas usa ferramentas diferentes, mas tem como base a interação entre pessoas. Atualmente, elas tomaram aspectos diversos para atender as demandas variadas de usabilidade que os usuários desejam realizar, o que as torna parte intrínseca do cotidiano das pessoas (SADAGHEYANI; TATARI, 2020).



Pela grande interatividade que as mídias conseguem oferecer, muitos pesquisadores desenvolveram técnicas eficazes para encontrar e estudar tanto grupos de pessoas como os interesses que guiam suas atividades nas redes (POCHAMPALLY; VARMA, 2011). Na literatura, conseguimos encontrar alguns exemplos de pesquisadores que usaram dados provenientes de redes sociais para avaliar a experiência de usuários e também tomar decisões baseadas nessas informações. Jiang et al (2021) realizaram um estudo de caso sobre a variação de sentimentos de turistas conforme a localidade em Pequim. Com este estudo, os autores propuseram um framework analítico para explorar e avaliar esses padrões de sentimentos de usuários por meio de dados de redes sociais (JIANG et al., 2021).

De maneira mais concentrada em produtos, Gozuacik, Sakar e Ozcan (2021) propuseram um framework para revelar as razões pelas quais um produto inovador pode falhar sob a perspectiva dos clientes e sugerir novas ideias para o desenvolvimento de produtos. Neste estudo, os autores realizaram uma série de estudos de modelos para serem avaliados nas redes sociais; como um estudo de caso, eles avaliaram o Google Class pelo Twitter (GOZUACIK; SAKAR; OZCAN, 2021). Como exemplo desta categoria de estudo na área de pagamentos mobile, Teng e Khong (2021) revelam que incrementaram uma metodologia que pegaram de outros pesquisadores com técnicas de estudo de usuários com dados de redes sociais; neste caso, eles avaliaram a experiência do usuário de e-wallets, uma ferramenta de pagamentos mobile. Os autores concluíram estatisticamente que a qualidade do sistema, da informação e do serviço têm impacto na satisfação e no engajamento do cliente. Além disso, eles identificaram características chaves do sistema de e-wallets (TENG; KHONG, 2021).

### **2.3. Twitter como fonte de dados**

Em meio à vasta gama de redes sociais, existem algumas com maior foco em compartilhamento de textos entre os usuários. O Twitter, que se encaixa dentro dessa categoria, é uma das redes sociais mais famosas entre os usuários e o volume de interações têm crescido exponencialmente (PHAN et al., 2021). De acordo com algumas estatísticas, no primeiro quarto de 2018 o Twitter contava com 330 milhões de usuários pelo mundo e, em 17 de abril de 2019, a rede gerava uma média de 6000 tweets por segundo (NOYES, 2021). Outrossim, na última década, houve um grande aumento no interesse de pesquisadores por usar redes sociais de microblogging, principalmente o Twitter, pelas suas características que favorecem a expressão de opiniões, bem como o manuseio e a exploração de textos (OUERTATANI; GASMI; LATIRI, 2021).

O Twitter possui uma variada gama de aspectos que diminuem a complexidade na compreensão de textos e atraem pesquisadores, como: limite máximo de caracteres (280); possibilidade de marcar o perfil de pessoas ou instituições ou responder conversas; a funcionalidade criar ou acompanhar discussões por meio de hashtags; e, por fim, seguir perfis (POCHAMPALLY; VARMA, 2011). Este conjunto de funcionalidades atrai estudos porque oferecem praticidade no tratamento dos dados, bem como medidas concretas sobre os interesses e conexões dos usuários da rede, que contém um grande acervo de clientes (PLATANIA; SPADONI, 2018). Ademais, o Twitter não é uma rede social restrita a um segmento de relações, como o LinkedIn que se restringe a conteúdos e interações do mercado de trabalho.

Nos últimos anos, o Twitter se tornou uma fonte notável e sólida de informações sobre a percepção que os usuários e stakeholders têm a respeito de uma determinada empresa, porque ali eles conseguem se expressar de maneira aberta e direta sobre os serviços e diretrizes das companhias (BARCHIESI; FRONZETTI COLLADON, 2021). Do mesmo modo, é viável obter uma extração significativamente precisa deste conteúdo nesta rede, já alguns estudos chegaram a correlacionar movimentos de mercados de capitais por meio da captação da volatilidade dos sentimentos dos usuários da rede (BOLLEN; MAO; ZENG, 2011). Esta

extração é feita a partir de mineração de textos e uso das indexações que a rede permite, métodos computacionais que já foram validados para avaliar emoções e até percepções de usuários a respeito de suas experiências com algumas empresas, seja com produtos ou serviços (PHAN et al., 2021).

Na literatura, conseguimos encontrar alguns exemplos de estudos que usaram o Twitter como fonte de dados. Fahmi, Hartoyo e Zulbainarni (2020) conduziram uma pesquisa no sentido de avaliar a imagem corporativa que os usuários possuíam a respeito da maior multinacional de mineração da Indonésia. Neste estudo eles levantaram uma série de variáveis que seriam analisadas em interações do Twitter, por meio de modelos de mineração de opiniões, para medir a imagem que a empresa possuía (FAHMI; HARTOYO; ZULBAINARNI, 2021). Ray, Bala e Jain (2020) também conduziram um estudo a fim de utilizar notas de emoção para classificar performances e prever avaliações. Os autores destacam a importância de entender as opiniões dos clientes para ter uma melhor estratégia prospectiva. Com o estudo, foi possível obter uma precisão significativa para antever as avaliações dos clientes sobre as empresas (RAY; BALA; JAIN, 2021). Outras pesquisas também mostram como é possível analisar dados do Twitter para identificar tendências sobre o comportamento dos usuários, além de possibilitar que análise seja aprofundada de maneira geográfica (PINDADO; BARRENA, 2020).

Também há pesquisas no Twitter para o setor bancário em particular. Estudos indicam que no segmento dos bancos, o uso de redes sociais é vital tanto para entender a visão que os clientes possuem a respeito de cada empresa, bem como para estabelecer relações que antes eram construídas nas agências, e agora migram para o relacionamento virtual (GANDOLFO, 2020). Abras e Matos (2020) desenvolveram um estudo de caso com dados de redes sociais, inclusive o Twitter, sobre a relação de um banco no Brasil com mais de 50 milhões de clientes. Neste estudo, os autores concluíram que o uso das redes sociais não exclui as demais maneiras de se comunicar com o cliente, mas são fundamentais para conhecer e gerar engajamento nos usuários a respeito dos produtos, bem como induzir ainda mais pessoas a se tornarem clientes, se possível antes que elas se tornem clientes de outros (ABRAS; MATTOS, 2021).

## **2.4. Mineração e análise textual em redes sociais**

Mineração de emoções em textos é um tipo de análise linguística feita por algoritmos computacionais que utilizam machine learning (ML) para avaliar, de acordo com parâmetros, uma grande base de dados (KIM et al., 2021). Esta análise tem por objetivo identificar as intenções de um determinado indivíduo, ou usuário, a respeito de um objeto (SERRANO-GUERRERO; ROMERO; OLIVAS, 2021). A análise de sentimentos tem sido um campo bastante estudado nos últimos anos, já que é uma área que consegue trabalhar com uma grande capacidade de dados e mostrar de maneira efetiva a opinião de um grupo por meio de modelos computacionais (ALI et al., 2021). A mineração de dados usa uma série de etapas para categorizar, filtrar e classificar as informações por meio de comparações com bases de dados previamente estudadas, que evoluem com o tempo por modelos de ML (DUONG; NGUYEN-THI, 2021). Desta forma, os modelos estão em constante evolução. Os modelos de ML dependem da qualidade dos inputs previamente selecionados, das bases de testes e de como as variáveis serão trabalhadas ao longo da análise para gerar aprendizado para o modelo, além da maneira de como o output será escolhido (BIRJALI; KASRI; BENI-HSSANE, 2021).

Capturar a opinião de um grupo de pessoas a respeito de política, produtos e eventos sociais é importante no processo de decisão de diversas instituições, que utilizam a mineração de opiniões para lidar com uma base tão grande de dados (ABDULLAH; RUSLI, 2021). Para compilar estas informações, é necessário possuir um canal de interações com o público desejado. As empresas buscam, então, nas redes sociais este acervo necessário para avaliar

as opiniões de seus usuários, uma vez que estes canais abrangem de maneira organizada e vasta as expressões das pessoas (MANOJ KUMAR et al., 2019). As redes sociais funcionam, pois, como um banco de dados eficaz para que as empresas colem e analisem as informações, interpretadas por algoritmos, a respeito do seu público (WELEKAR; KARANDIKAR; TIRPUDE, 2020).

Há vários estudos que utilizam modelos de mineração de opiniões em áreas diversas. Na seção 2.3 foram citados estudos que utilizaram o Twitter como fonte de dados para as análises. Em grande parte, foi usado mineração de textos para se obter o resultado final. Outros estudos também foram feitos no sentido não só de coletar e avaliar as experiências, mas de obter ferramentas que pudessem prever as atitudes dos clientes em relação às marcas (RAY; BALA; RANA, 2021). De maneira mais específica, o setor financeiro oferece uma gama extremamente relevante de oportunidades para se antecipar tendências e também compreender as mudanças de comportamento assim que acontecem, já que os serviços bancários são extremamente complexos (GUPTA et al., 2020).

A seguir são apresentadas as etapas da metodologia que será usada para realizar a pesquisa. Na seção de metodologia, será descrito o processo usado para selecionar materiais para revisão bibliográfica. Será detalhado o planejamento para a execução da mineração de texto, desde a coleta até o resultado final da análise textual; este tipo de análise de dados possui etapas gerais, previamente estabelecidos por pesquisadores, que serão detalhados durante a seção de metodologia. E, por fim, será abordado o caminho percorrido para executar a análise dos dados.

### **3. METODOLOGIA**

O caminho seguido para a pesquisa foi consolidado a partir de um framework proposto por Stieglitz et al. (2018) o qual é estruturado nas seguintes partes:

1. Descoberta (Discovery): esta é a etapa onde os tópicos a serem analisados serão extraídos. Para isso, definem-se as palavras chaves, indexadores e temas que abrangem os tópicos selecionados.
2. Extração ou rastreamento (Tracking): esta parte se baseia em decidir as fontes de dados (Twitter), abordagem, método, e saída de dados.
3. Preparação (Preparation): esta etapa tem como objetivo a alocação e preparação dos dados extraídos para análise futura, como limpeza de textos, e remoção de dados duplicados ou sem relevância.
4. Análise (Analysis): a fase de análise de dados é constituída de diversos métodos disponíveis a partir da quantificação dos dados coletados com relação a um algoritmo de aprendizagem de máquina (ML).

Este framework foi executado em três principais etapas: revisão da literatura que abrange o primeiro passo (Discovery), coleta e processamento de dados que abrange os passos 2 e 3 (Tracking and Preparation), e análise dos resultados que abrange o passo 4 (Analysis). Desta maneira, o framework desenvolvido busca estruturar etapas fundamentais em uma análise de dados de redes sociais (STIEGLITZ et al., 2018). Estas três etapas se relacionam como mostra a figura 1:

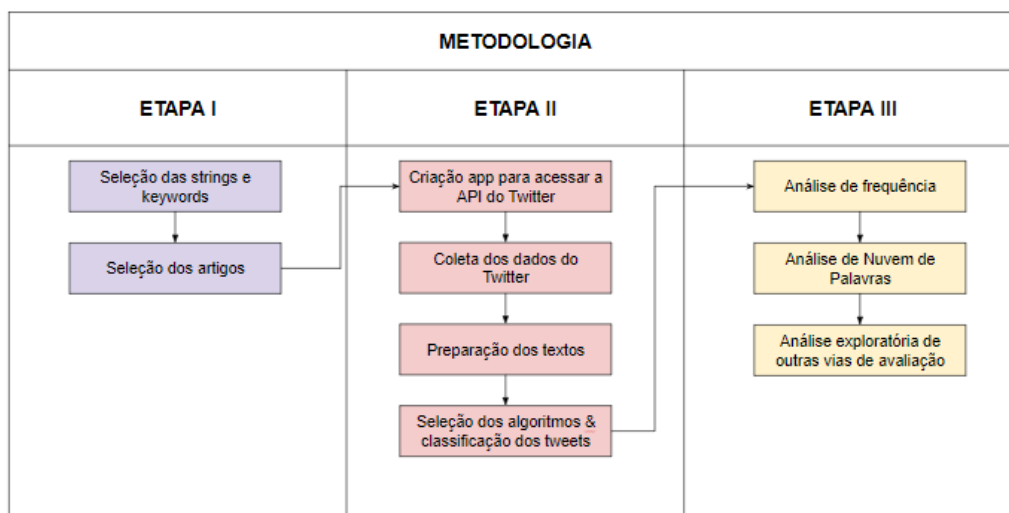


Figura 1 - Etapas da Metodologia (Fonte: Autor, 2021)

### 3.1. Etapa I – Revisão da literatura

A primeira etapa buscou reunir o conhecimento necessário para que fosse compreendido o contexto dos objetos analisados, as possibilidades de análise das ferramentas selecionadas e os fundamentos de experiência do cliente. A busca dos artigos foi realizada com o intuito de produzir o conhecimento necessário para que a pesquisa pudesse ser realizada. Foi realizada uma revisão da literatura não exaustiva, pois o intuito era apenas fundamentar os conceitos abordados. Segundo Webster e Watson (2005), a revisão bibliográfica é essencial para qualquer trabalho acadêmico. Uma revisão eficaz cria uma base para o desenvolvimento do conhecimento, já que ela auxilia na identificação de campos de estudo consolidados e de áreas com ausência de conteúdo; a fim de que se explore o que foi construído para o avanço do estudo (WEBSTER; WATSON, 2005).

A fim de realizar a busca de artigos para a revisão da literatura, foi usada a base de dados Scopus, da Elsevier e a busca dos artigos foi feita apenas em inglês. A seleção dos artigos passou por um critério específico de data de publicação. Foram coletados das bases apenas artigos publicados nos últimos cinco anos, a partir de 2017, enquanto os demais foram selecionados a partir destes trabalhos ou pelo conhecimento prévio do conteúdo pelo autor. A busca de artigos passou por cinco temas principais. Para todos os temas, a busca na base de dados foi feita a partir da seleção das strings contidas nos títulos, resumos e palavras-chaves. Vale ressaltar que nem todos os artigos selecionados foram diretamente referenciados no trabalho, mas mesmo assim exerceram alguma influência sobre a construção teórica da pesquisa. Os temas:

- *Fintech* – Strings: “*fintech*” e “*innovation*”. No total, 11 artigos foram selecionados.
- User Experience (UX) / User Centered Design (UCD) – Strings: “*user experience*”. Ao final, 17 artigos foram selecionados
- Mídias Sociais – Strings: “*Social Media*”. Ao todo, foram selecionados 13 artigos.
- Twitter – Strings: “*Twitter*” e “*Text Mining*”. Ao final, 12 artigos foram selecionados.
- Mineração de textos – Strings: “*Text mining*”. Foram selecionados 11 artigos.
- Outros tópicos – 8 artigos.

Além da revisão da literatura, foi realizada uma revisão técnica das ferramentas que foram utilizadas ao longo da pesquisa. Para isso, utilizamos o curso “Mineração de emoção em textos com Python e NLTK”, na plataforma Udemy. Este curso forneceu uma base teórica

sólida para que fosse possível entender a aplicação dos conceitos e os fundamentos da análise feita por meio da mineração textual e análise de sentimentos.

Esta etapa teve extrema importância para construir uma base de conhecimento a respeito das possibilidades que a mineração de textos oferece, e como aliá-la ao estudo de experiência do cliente em um segmento tão inovador como o das *fintechs*. Desta forma, este estudo viabilizou o entendimento do que poderíamos extrair com a coleta de dados e nos auxiliou na interpretação dos resultados obtidos. Assim, conseguimos detalhar as partes de execução do estudo.

## 3.2. Etapa II – Coleta e processamento dos dados

### 3.2.1. Coleta dos Dados

A coleta dos dados deu início à parte 2 do framework apresentado. Com ela, foi executado o planejamento para obter os dados que seriam analisados. Para coletar os dados do Twitter, foi necessário criar uma conta de desenvolvedor. Com esta conta, foi possível criar uma “application” (app) para obter acesso à API do Twitter. Nesta API, a rede disponibiliza, de acordo com a app criada, os tweets que poderão ser coletados para posterior análise. Quando se cria uma app, o Twitter cria chaves de acesso únicas que serão usadas para que o Twitter libere acesso e assim consiga controlar como os usuários estão coletando dados da sua plataforma. Cada app pode estar sujeito a limitações de volume e categorias de informações (TWITTER INC., 2021a). Com essas chaves geradas, o usuário pode utilizar programas na linguagem de programação Python que façam uma busca ordenada pela API do Twitter com alguns filtros selecionados. A própria rede disponibiliza uma biblioteca em Python para que o usuário consiga navegar em sua rede para obter os dados desejados (TWITTER INC., 2021b).

Desta maneira, para acessar a API do Twitter, foi construído o código exibido na figura 2:

```
import tweepy
import pandas as pd

#credenciais da API do twitter
c_key = "chave-pessoal"
c_secret = "chave-pessoal"
a_token = "chave-pessoal"
a_token_secret = "chave-pessoal"

#inicializando nossa conexão com a API do twitter
auth = tweepy.OAuthHandler(c_key, c_secret)
auth.set_access_token(a_token, a_token_secret)
api = tweepy.API(auth, wait_on_rate_limit=True)
```

Figura 2 - Código de acesso (Fonte: Autor, 2021)

A função “api” (última linha) foi criada para ser utilizada sempre que for necessário acessar os dados da API.

#### 3.2.1.1. Escolha das empresas a serem comparadas

Neste estudo, optou-se por fazer a comparação entre empresas, e não modelos, para simplificar a coleta de dados, uma vez que os usuários tendem a se referir às empresas pelo nome. Optamos por escolher apenas uma de cada ramo para que os resultados pudessem ser analisados sem as nuances que as empresas podem ter dentro de um mesmo ramo. Para isso, escolhemos duas empresas, uma para cada ramo estudado: os bancos tradicionais e as *fintechs*. Nosso critério de decisão foi estritamente objetivo e numérico, escolhemos a maior empresa privada de cada ramo no Brasil. Desta forma, as duas empresas selecionadas

foram Nubank (Nu Pagamentos SA), maior *fintech*, e Banco Bradesco, maior banco privado (BACEN, 2021). Vale ressaltar que por *fintech* buscamos nesta pesquisa aquelas que ofereçam pelo menos os produtos cartão de crédito e conta corrente para que a comparação pudesse ser feita no mesmo escopo de usabilidade de um banco de varejo. As empresas que oferecem serviços financeiros que não englobam estes produtos, ainda que fossem maiores que o Nubank, foram descartadas.

### 3.2.1.2. Extração dos Tweets

Para a análise da pesquisa, foram utilizados critérios com objetivo de evitar vieses na fonte de dados (etapa 1 do framework):

- Os tweets deveriam conter o nome do banco
- Os tweets não poderiam conter a palavra “limite” - pois ela diz respeito ao produto do banco e não à experiência de como ele é oferecido
- Os tweets não poderiam ser um retweet (quando algum usuário apenas compartilha o que outro postou) – considera-se que um retweet não tem a mesma intensidade na expressão do usuário e, assim, por termos uma amostra não exaustiva, é mais proveitoso ter homogeneidade neste aspecto.
- Os tweets deveriam ser postados em português
  - porque foram usadas empresas que tem visibilidade global, porém o foco da pesquisa é o público nacional;
  - como o filtro mais direto de país exclui mais usuários, que não dão acesso à sua localidade, o filtro da língua pode ser uma aproximação razoável neste caso.
- Os tweets não poderiam ser postados pela própria empresa
- Por fim, foram excluídos tweets que continham a palavra doença por coletarmos muitos tweets sobre ações solidárias, onde o nome do banco era apenas a ferramenta para coletar o dinheiro.

Nesta pesquisa em específico, definimos que a nossa amostra precisaria ser a partir de 1000 tweets. Usamos o código de busca, como demonstra a figura 3, por algumas iterações até conseguirmos um número maior que 1000 tweets distintos. Com esse parâmetro, nossa busca se encerrou em 1088 tweets para a *fintech*. Fizemos em paralelo um ciclo idêntico para o banco e ajustamos o tamanho da base para que ambas as empresas contassem com uma amostra de mesmo tamanho para a análise. Estes dados foram coletados durante os meses de novembro de 2021, dezembro de 2021 e janeiro de 2022 de maneira aleatória com relação à data, dentro do intervalo selecionado, e também aleatória com relação aos perfis ou regiões de emissão do tweet. Esta quantidade de tweets, 1088, é o número final de textos após excluirmos os tweets repetidos que foram coletados durante os ciclos.

Para fazer a busca dos tweets segundo as instruções, foi utilizado o código da figura 3:

```
1 # construir a query (algoritmo) de busca a ser usado na API do Twitter
2 query = "(NUBANK OR Nubank OR nubank) lang:pt -limite -doença -filter:retweets -from:Nubank"
3 busca = api.search(query,
4                  count=100,
5                  tweet_mode="extended")
```

Figura 3 - Código de busca (Fonte: Autor, 2021)

Vale ressaltar que, pelo código da figura 3, fazemos uma busca com um número limitado de tweets por conta de uma limitação na própria API do Twitter que restringe a busca por quantidade. Ainda que a rede não fizesse esta restrição, o possível volume de dados poderia exigir muito da capacidade de processamento, logo optou-se por realizar apenas uma busca não exaustiva de tweets em relação ao tema. Para checar a condição inicial dos dados,

separou-se uma amostra de 15 tweets para serem lidos a fim de entender se os textos eram pertinentes aos objetos estudados. Nesta amostra, todos os textos eram pertinentes e mostrava expressões do usuário emissor em relação ao objeto da busca.

Uma vertente que buscamos abordar neste trabalho era a análise por distribuição geográfica. Nesta vertente, iríamos entender se existe alguma variação na experiência dos usuários que possa se relacionar com a região em que o usuário se encontra influenciada pelos costumes ou serviços que uma determinada região pode oferecer. Porém, entre os dados que coletamos, verificamos que a maioria não possuía os dados geográficos disponíveis. Por conta disso, esta análise ficou inviável.

### 3.2.2. Preparação dos textos

A preparação dos textos deu início à parte 3 do framework. Com o acesso à API do Twitter, é possível coletar dados das mais diversas formas, obedecendo as restrições que a rede coloca a cada app. Ao coletarmos informações de alguma rede social, nos deparamos com a situação de que os textos não têm uma padronização formal, contêm subjetividades e apresentam variações aleatórias com relação a região que se encontram bem como ao perfil do usuário que publicou aquela informação. Para isso, é importante fazer uma preparação. Após a coleta dos textos, eles foram passados para uma tabela, com todas as informações que o Twitter disponibiliza sobre cada tweet. Em seguida, para cada tweet, foi feita uma preparação para que os textos pudessem ser interpretados pelo algoritmo selecionado.

O algoritmo de classificação de textos apenas consegue interpretar textos processados de uma maneira padronizada. Quando os textos são coletados da rede social, em seu estado original, eles estão em um formato diferente que contém vários empecilhos para a devida interpretação do conteúdo. A preparação é feita em três etapas.

A primeira etapa consistiu em tirar caracteres de indexação, links e pontuação. Esta parte do texto não oferece nenhum valor analítico para entender qual é o sentimento do usuário em relação ao objeto. A segunda etapa foi remover as palavras que tem baixa relevância analítica como preposições ou artigos, uma vez que o algoritmo leva em consideração a frequência das palavras e poderia ser poluído com palavras que não adicionam valor à análise. A terceira etapa foi passar cada palavra do texto para o seu radical. Isto é usado para que não houvesse distinção entre tempos verbais, gênero ou quantidade. Assim, palavras que indicavam a mesma ideia tiveram o mesmo valor (etapa 3 do framework).

As funções que realizam a preparação dos textos são, respectivamente, como exibido na figura 4:

```

1 # Função para limpar o texto que será analisado - aqui retiramos os links, #, @ e o nome do banco em estudo
2
3 def limpeza_dados(texto):
4     # remove links, pontos, ponto e vírgula dos tweets
5     texto = re.sub(r"http\S+", "", texto).lower().replace(".", "").replace(",", "").replace("-", "")
6     .replace(":", "").replace(";", "").replace("(", "").replace(")", "").replace("&", "")
7     # texto = re.sub(r"[^a-zA-Zâ-ûø-9]", "", texto.lower()) #limpa e converte para minúsculo
8     texto = re.sub(r"nubank", "", texto.lower()) # remover a palavra nubank para que não entre no cálculo
9     texto = re.sub(r"#", "", texto.lower()) #remover as hashtags
10    texto = re.sub(r"@", "", texto.lower()) #remove os @s
11    return(texto)

```

```

1 # função para retirar as palavras que não possuem relevância para a análise
2
3 def RemoveStopWords(texto):
4     stopwords = set(nltk.corpus.stopwords.words('portuguese'))
5     palavras = [i for i in texto.split() if not i in stopwords]
6     return " ".join(palavras)

```

```

1 # Função de stemming das palavras em cada tweet
2
3 def stemming(texto):
4     teststemming = []
5     stemmer = nltk.stem.RSLPStemmer()
6     for palavras in texto.split():
7         comstem = [p for p in palavras.split()]
8         teststemming.append(str(stemmer.stem(comstem[0])))
9     return teststemming

```

Figura 4 - Código de preparação de textos (Fonte: Autor, 2021)

### 3.2.3. Seleção De Algoritmos

Como algoritmo de interpretação e análise de textos, usamos a metodologia Natural Language Toolkit (NLTK) para avaliar os textos de acordo com uma base textual padronizada para comparação. Este algoritmo tem diversas funcionalidades que podem nos mostrar informações relevantes para julgar a qualidade da nossa análise como acuracidade do modelo, palavras mais relevantes e a distribuição da polaridade da amostra (NLTK, 2021).

Dentro da biblioteca NLTK, podemos encontrar uma subparte chamada Classificador Naive Bayes. Este classificador é responsável por realizar os cálculos necessários de distribuição das palavras que definem a polaridade de um determinado texto. Com a classificação previamente feita de alguma base confiável, o algoritmo Naive Bayes consegue gerar os padrões para que os demais textos sejam classificados, e por meio do cálculo que o algoritmo realiza, podemos ter a classificação com uma nota que revela o quanto cada amostra têm de uma determinada polaridade (NLTK, 2021).

#### 3.2.3.1. Algoritmo de classificação de Frases Coletadas

Para analisar os textos com o Naive Bayes, foi preciso usar uma base de textos já preparada e classificada para que fosse possível comparar os dados coletados com esta base, que o algoritmo toma como referência de textos positivos ou negativos.

Desta forma, juntamos algumas frases de bibliotecas que são comumente usadas como referência e adicionamos algumas classificações manuais de tweets dentre os que foram coletados.

#### 3.2.3.2. Análise de sentimentos

Com esta base preparada, fizemos o treinamento do algoritmo e com ele criamos a função usada para classificar cada texto. Assim, criamos uma coluna na tabela em que registramos os textos para registrar também a classificação que o algoritmo deu a cada texto. O algoritmo dá a cada texto uma nota de 0 a 1, onde 0 é o máximo de negatividade e 1 é o máximo de positividade. Para este estudo em específico, consideramos que os textos seriam



classificados como positivos apenas com uma nota mínima de 0.6 e negativos com uma nota máxima de 0.4. Os textos que ficassem entre 0.4 e 0.6 seriam considerados neutros.

Para isso, utilizamos o código da figura 5 para construir o classificador, onde a “basecompleta” é a base já processada que serviu de comparação com os textos:

```
1 # constroi o classificador tomando como parâmetro a base de treinamento agora preparada
2 classificador = nltk.NaiveBayesClassifier.train(basecompleta)
```

Figura 5 - Código de construção do classificador (Fonte: Autor, 2021)

E depois, utilizamos o código da figura 6 para montar uma função que classificasse, dentro do *dataframe*, cada um dos textos coletados:

```
1 # função que pega a nota do classificador para cada texto que for passado
2
3 def valor_nota_alegria(texto):
4     distribuicao = classificador.prob_classify(texto)
5     count = 0
6     for classe in distribuicao.samples():
7
8         count += 1
9         if count > 1:
10             break
11
12     x = str("%s: %f" % (classe, distribuicao.prob(classe)))
13     y = x.split()
14     z = float(y[1])
15     return z
```

Figura 6 - Código de atribuição de notas aos textos (Fonte: Autor, 2021)

### 3.3. Etapa III – Análise dos resultados

#### 3.3.1. Análise de sentimentos

Com a etapa três, deu-se início ao último passo do framework. Nesta seção, o objetivo é quantificar a polaridade dos dados coletados. Assim, usamos as saídas que o algoritmo teve por meio do classificador para que pudéssemos interpretar os textos.

##### 3.3.1.1. Análise de frequência

Após a classificação de todos os textos, fez-se uma contagem de quantos textos eram positivos, negativos ou neutros dentro da amostra obtida para estudo. Neste estudo em específico, fazemos o cálculo tomando como o base os 1088 tweets selecionados sobre cada uma das financeiras de acordo com as diretrizes já apresentadas.

##### 3.3.1.2. Nuvem de Palavras

Para complementar a análise dos textos, separou-se os tweets para que fosse possível realizar a análise da frequência que as palavras aparecem no total dos dados coletados. Esta análise é visualizada com um algoritmo chamado “Nuvem de Palavras”, onde o tamanho com que a palavra aparece reflete a frequência com que ela ocorre nos dados coletados.

Desta forma, separamos para cada instituição os tweets positivos e negativos para que pudéssemos identificar se a frequência de uma determinada palavra mostra satisfação ou insatisfação. Da mesma maneira, a separação das polaridades nos ajuda a direcionar possíveis insights a partir do que os algoritmos nos mostrarem.

Para esta parte da análise, nós utilizamos os textos já preparados, uma vez que aqui também desejamos excluir palavras que geram pouco valor analítico (Por exemplo: “de, da, e”). Para isso, usamos colunas diferentes da preparação até o nível de tirar palavras de baixo valor analítico e a preparação que deixa apenas os radicais das palavras para entrarem no algoritmo de classificação.

Para construir este diagrama chamado “Nuvem de palavras”, utilizamos o código da figura 7:

```
1 import matplotlib.pyplot as plt
2 plt.style.use('fivethirtyeight')

1 from wordcloud import WordCloud

1 # fazer a nuvem de palavras com o dataset limpando o texto e tirando stopwords
2 # Word Cloud
3 todas_as_palavras = ' '.join( [twts for twts in df['text']] )
4 wordCloud = WordCloud(width = 500, height = 300, random_state = 21, max_font_size = 119).generate(todas_as_palavras)
5
6 plt.imshow(wordCloud, interpolation = "bilinear")
7 plt.axis('off')
8 plt.show()
```

Figura 7 - Código de construção da nuvem de palavras (Fonte: Autor, 2021)

#### 3.3.1.3. Análise exploratória da experiência dos usuários dos bancos selecionados

Para completar os resultados que a análise de sentimentos trouxe, fizemos uma análise sobre outras avaliações de satisfação do cliente que são difundidas no mercado junto a uma análise de engajamento das duas marcas selecionadas com dados da busca pelas empresas. Desta maneira, conseguimos comparar se os resultados indicam consistência com o que vemos em outras vias de avaliação bem como contextualizar os insights encontrados ao longo da análise.

Para contextualizar estes resultados, optou-se por compará-los com algumas vias mais formalizadas de avaliação da experiência do consumidor. Todos os dados das vias usadas na análise exploratória foram coletados no dia 30 de dezembro de 2021. Para cada via, temos os seguintes objetivos e pontos de atenção:

- Lojas de APPs
  - Objetivo: Como todos os consumidores precisam necessariamente utilizar uma das duas lojas, Google Play ou App Store da Apple, para acessar o aplicativo, as lojas podem apresentar uma avaliação mais representativa e confiável da experiência
  - Observação: Como o intuito da loja é fornecer o aplicativo apenas, a avaliação pode enviesar o cliente a avaliar na plataforma apenas o funcionamento e usabilidade que o aplicativo oferece, sem considerar outras dimensões da sua experiência com a empresa
- Reclame Aqui
  - Objetivo: A plataforma da Reclame Aqui é focada em intermediar a notificação e solução de problemas que os clientes podem ter com empresas. Assim, podemos avaliar o quanto uma determinada empresa gera problemas graves o suficiente para gerar uma reclamação formal e o quão bem ela consegue solucionar essas questões
  - Observação: Esta plataforma, entretanto, tem foco em intermediar conflitos entre clientes e empresas. Por isso, esta plataforma não traz tanta visibilidade de boas experiências quanto traz de experiências ruins. Assim, ela não oferece tantas informações sobre quão boa uma experiência pode ser, e também não auxilia a separá-las de experiências intermediárias.

- NPS
  - Objetivo: A pesquisa de NPS (Net Promoter Score) contempla um escopo um pouco mais amplo do que as demais. Ela direciona o cliente a avaliar toda a visão que ele tem da experiência com a marca levando em conta as diversas interações que pode ter tido com ela, desde o uso dos produtos até o relacionamento nas redes sociais (SOLUCX, 2021). A pesquisa avaliou 9 empresas do ramo bancário, entre elas as duas selecionadas para este estudo.
  - Observação: A pesquisa de NPS teve duas restrições básicas:
    - Ela utilizou uma população restrita a residentes de São Paulo. Isto pode gerar enviesamento em relação às nuances que as diferentes regiões podem apresentar
    - Ela oferece apenas uma nota que os clientes dão com relação à percepção que têm da empresa. Cada nota pode estar atrelada a diferentes motivos e, logo, dificulta a busca pelos principais pontos de insatisfação dos clientes.
- Google Trends
  - Objetivo: Esta ferramenta do Google oferece a possibilidade de comparar a busca que cada empresa teve em sua plataforma ao longo do tempo e entre as diversas regiões do país. Esta plataforma foi usada para se entender qual foi a tendência de ambas as empresas com relação às buscas no Google e qual possuiu um volume maior.
  - Observação: Esta ferramenta mostra apenas a busca por objetos que possam envolver o nome do banco em questão. Nestas buscas não mostram se os usuários são ou desejam se tornar clientes do banco. A busca também não mostra se o interesse é positivo ou negativo, mas apenas o número de vezes que os objetos da busca continham o nome do banco.

#### 4. Resultados

Os resultados foram divididos em duas seções com objetivos complementares. A primeira parte consistiu dos resultados da interpretação dos dados coletados no Twitter. Esta foi dividida em duas subseções: análise de frequência e nuvem de palavras. A primeira subseção traz a frequência com que os textos têm um sentimento negativo, positivo ou neutro em relação ao objeto. Com isso, podemos comparar a imagem que os usuários têm de cada empresa, uma vez que a empresa cujos usuários se expressam de maneira mais positiva tende a ter um relacionamento melhor com seus clientes. Já a segunda subseção consistiu em analisar quais eram as palavras mais frequentes de acordo com o sentimento daquele texto já interpretado pelo algoritmo. Com isso, buscou-se entender qual era a relação dos temas e das características que influenciam o usuário a ter cada um dos sentimentos.

A segunda parte consistiu numa análise exploratória de engajamento e da expressão dos clientes sobre a experiência em cada empresa. Esta análise foi realizada a partir de algumas vias consolidadas que fazem o intermédio dos clientes com as empresas. Assim, esta parte foi dividida nas vias estudadas. Desta maneira, cada via traz uma ótica distinta a respeito do engajamento ou da imagem que o cliente pode ter com a empresa. Esta parte inicia com as lojas de aplicativos, via que traz a avaliação dos usuários sobre os aspectos técnicos dos produtos digitais que as empresas oferecem. Em seguida, estudou-se o volume de busca que cada empresa possui na plataforma do Google como um indicativo do alcance de cada empresa. Então, analisou-se como cada empresa lida com situações problemáticas com seus clientes a partir das informações coletadas pela plataforma do Reclame Aqui. Na sequência,

utilizou-se as informações da pesquisa de NPS, que consiste em avaliar o quanto um cliente promoveria uma determinada empresa. Com estas informações, é possível avaliar critérios mais subjetivos em relação à satisfação do cliente durante a experiência. Ao fim, analisou-se o engajamento que cada empresa possui em duas redes sociais consolidadas, Twitter e Instagram, a fim de entender o quanto cada empresa consegue atrair clientes para um relacionamento de lealdade com a marca. Desta forma, o panorama dado pelas duas partes de estudo busca trazer qual empresa proveu uma melhor experiência sobre para os clientes sobre diversos aspectos.

#### 4.1. Análise de Sentimentos dos dados coletados do Twitter

Na análise de sentimentos, o objetivo foi pegar dados disponibilizados pela rede social Twitter em sua forma original e processá-los por meio de um algoritmo que pudesse interpretar um alto volume de informações. Com este processo, obtivemos uma classificação dos textos com relação à polaridade que eles tinham com relação ao objeto em questão. A saída que o algoritmo traz é uma nota onde o mínimo é 0, completamente negativo, e o máximo é 1, completamente positivo. Assim, seguimos para extrair uma análise de frequência e a nuvem de palavras.

##### 4.1.1. Análise de frequência

Após a classificação de todos os textos coletados, obtivemos as frequências como exibido na figura 8. Nesta parte, levamos em consideração uma contagem direta e mostramos os resultados em porcentagem, onde o todo é formado por todos os tweets coletados separados por empresa.

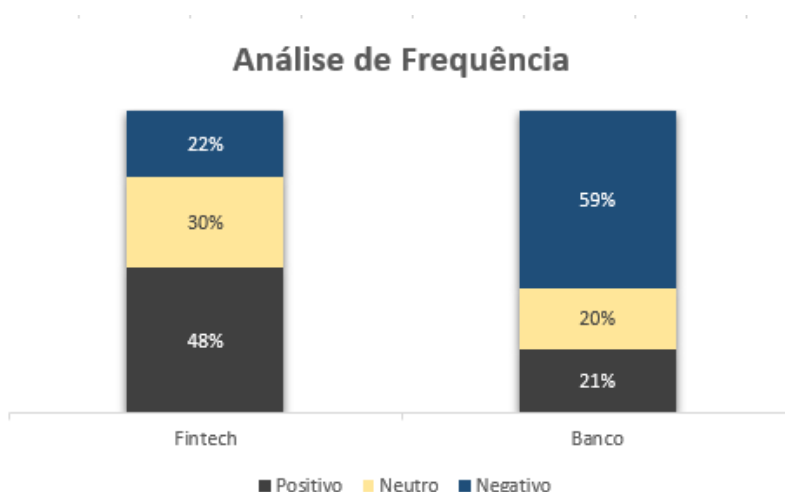


Figura 8 - Gráfico de frequência dos sentimentos nos textos (Fonte: Autor, 2022)

Ao quantificar a frequência de sentimentos positivos, negativos e neutros, vemos que a *fintech* possui uma relação de 48% de sentimentos positivos, 30% de neutros e 22% de negativos. Nesta mesma análise, o banco teve uma relação de 21% de positivos, 20% de neutros e 59%

de negativos. Desta forma, conforme a figura 8, o banco tem mais da metade da amostra com expressões negativas com relação a *fintech*. Isto indica que o banco pode ter a imagem pública sendo detratada pelos usuários da rede. Neste mesmo ambiente, vemos que a *fintech* possui quatro quintos da amostra com expressões positivas ou neutras. Isto indica que a imagem pública da empresa pode estar sendo valorizada. Este fenômeno pode ser influenciado diretamente pela composição de clientes da base com relação aos que utilizam a rede social Twitter. Como a *fintech* é totalmente digital, ela tende a explorar melhor esta conexão e atrair pessoas com mais identificação por plataformas digitais como o Twitter. Enquanto isso o banco tradicional tende a não conseguir tanta identificação com os usuários ativos das mídias sociais com o seu modelo de negócio.

Outro fator que poderia influenciar esta diferença seria a distribuição de idade das bases. É provável que o banco digital tenha uma média de idade da base menor que o banco tradicional. Com esta situação, pessoas mais velhas e estabilizadas financeiramente tendem a ter uma maior exigência com relação aos produtos financeiros.

Vale notar também que, de qualquer forma, a empresa digital possui cerca de um quinto da amostra com expressões negativas. Isto poderia estar relacionado ao fato de que os usuários tendem a se expressar quanto algo não corriqueiro acontece. Caso isso seja verdadeiro, muitos usuários com boas experiências não iriam se expressar por considerar a boa relação com a empresa como rotineiro. De qualquer maneira, notamos que esta proporção é significativa na amostra.

Porém, é notável que existe uma vantagem para a *fintech* quanto à avaliação da experiência pela rede social Twitter. Como a rede social indica um lugar de maior espontaneidade e intimidade com pares, o melhor desempenho da *fintech* indica que ela tem uma melhor performance em construir confiança com os clientes, bem como tende a usufruir mais do efeito de atração que os clientes podem exercer com seu círculo de amizade com relação às empresas.

É importante salientar que não foi encontrada uma maneira confiável de avaliar a representatividade da base de dados coletada. A representatividade deveria ser comprovada sobre dois aspectos, a primeira seria da base de clientes em relação às pessoas que se expressam nas redes sociais, e a segunda da base de usuários em relação à amostra coletada para estudo. A primeira relação pode não ser simétrica uma vez que é provável que pessoas com mais idade não sejam tão engajadas nas redes sociais como pessoas mais novas. A segunda relação poderia ser assimétrica uma vez que não temos dados mais específicos sobre os usuários que postaram cada texto. Por isso, existe a possibilidade de o resultado estar enviesado e, por mais que ele consiga trazer uma visão sobre uma amostra considerável sobre a experiência destes clientes, ele pode não mostrar a real visão do público geral. Por esse motivo, existe um valor grande em buscar respaldo destes resultados em outras vias consolidadas no mercado.

#### 4.1.2. Nuvem de Palavras

Com o intuito de entender onde se concentram os principais tópicos que dão a vantagem à *fintech* em relação ao banco tradicional, buscamos realizar a análise da Nuvem de Palavras. Com ela, obtivemos os resultados para ambas as empresas como mostrado nas figuras 9 e 10, sobre o conjunto de textos positivos e negativos. A figura 9 mostra as palavras que são mais frequentes nos tweets relacionados com a *fintech*, onde o lado A mostra as palavras dos tweets classificados como positivos, enquanto o lado B exibe as palavras dos negativos.



Figura 9 - Nuvem de palavras sobre a *fintech* Nubank (Fonte: Autor, 2022)

A figura 10 mostra a mesma relação para os tweets relacionados ao banco, também com lado A para tweets positivos e B para tweets negativos.



Figura 10 - Nuvem de palavras sobre o banco Bradesco (Fonte: Autor, 2022)

Com a análise revelada pela figura 9, as nuvens da *fintech*, podemos perceber que os tópicos gerais sobre os quais os clientes discorrem são parecidos; tanto nos casos positivos, lado A, como negativos, lado B. Observa-se, em tamanhos variados, palavras como “cartão”, “banco”, “conta”, “fatura”. Isto nos mostra que os clientes estão comentando sobre o uso rotineiro dos produtos (algumas *fintechs* são coloquialmente chamadas de bancos). Porém, vemos o aparecimento de algumas palavras periféricas também muito similares ainda que apareçam em tamanhos diferentes quando comparamos o grupo positivo com o negativo. Isso indica que os clientes estão se expressando sobre experiências muito parecidas, porém com perspectivas diferentes. Esta variação na perspectiva pode ser originada por uma jornada semelhante do cliente dentro do serviço, porém com variações na qualidade da experiência. Estas variações tendem a ocorrer pela diferença de familiaridade que alguns usuários tem com produtos digitais.

Quando analisamos as nuvens do banco tradicional, tanto negativa quanto positiva, conseguimos enxergar também palavras como “banco”, “cartão” e “conta” que representam assuntos a respeito do produto oferecido. Porém, há também uma forte influência nos dados a respeito de alguns eventos que envolveram o banco selecionado no intervalo de coleta dos dados. Isso pode ser representado por palavras “propaganda”, “marketing”, “agro”, “pecuarista” que aludem aos eventos de uma propaganda de marketing que o banco realizou

com referências ao agronegócio. Desta maneira, seria necessário um intervalo maior de tempo para entender o quanto estes sentimentos poderiam variar. De qualquer forma, o resultado sugere que o banco tradicional teve um problema em se conectar com os usuários das redes sociais por meio de sua propaganda.

#### 4.1.3. Amostra de tweets classificados

Assim que os textos foram classificados, foi possível pegar alguns destes tweets para que eles fossem analisados de maneira mais próxima. Assim, pode-se gerar um valor adicional com respeito a maneiras de se olhar para as expressões do cliente. Esta parte do processo pode ser usada caso não haja a necessidade de se interpretar muitos textos, e objetivo seja apenas de classificar-los com o algoritmo e em seguida analisá-los mais cuidadosamente. Com isso, separamos uma amostra dentro dos tweets selecionados de cinco tweets positivos e cinco negativos de cada empresa que pudessem nos trazer de modo mais claro quais são as expressões dos usuários. Por isso, temos a seguinte amostra:

- Tweets Positivos Fintech

- a) “@nubank Já consegui resolver no app. Graças a Deus estornaram o valor e vão me enviar outro cartão. Obrigada pela rapidez de sempre”
- b) “Muito lindo tá @nubank #UltraVioleta (link do produto)”
- c) “Todos os meus planos pra 2022 começam com o saldo da mega da virada rendendo no nubank. Se não rolar, será mais um ano vivendo à sorte do destino.”
- d) “@acaodacidadania Grande iniciativa. Fazendo minha doação agora. Parabéns ao @nubank”
- e) “@crissilvaeng O comprovante do Nubank é o mais fácil, fica a dica...”

É possível notar que os usuários se expressam a respeito de diversos temas nos tweets positivos. O primeiro texto revela a satisfação recorrente de um usuário com o serviço prestado pelo aplicativo. O texto B revela a admiração do usuário pelo design de um dos produtos lançados pela fintech durante o ano de 2021. No texto C, o usuário demonstra, por meio de uma expressão em tom de humor, que tem preferência pela fintech; isto revela uma boa relação com os clientes, uma vez que o usuário se sente confortável de gerar interações com amigos e incluir a empresa nesta situação. O texto D revela admiração pelo posicionamento da fintech em auxiliar uma situação de crise e facilitar a participação dos clientes neste movimento. Por fim, no texto E, o usuário fala sobre a facilidade de utilizar o produto da fintech, o que indica um design do produto bem construído para a experiência do cliente.

- Tweets Negativos Fintech

- a) “E aí @nubank. Não consigo fazer pix. Não consigo pagar boleto.”
- b) “@nubank Mas nem na função de inserir funciona, consta como “senha bloqueada” mesmo sem eu ter errado a senha nenhuma vez”
- c) “@mimadodoceu @nubank Mas no site da lojinha tá lá o pedido, só não enviam”
- d) “Eu tô no desespero já com o @nubank, tenho um dinheiro lá que queria usar, mas não consigo abrir meu app para desbloquear meu cartão. Da um help”
- e) “@dyghu\_ @nubank Mano o meu dinheiro sumiu , e eles se fazendo de sonso”

Nos tweets negativos, nota-se que os clientes se expressam mais com relação ao produto ou serviço que a fintech oferece. Os textos A e B falam sobre a dificuldade dos usuários de



utilizarem o aplicativo para executarem determinadas atividades. No texto C, o usuário demonstra insatisfação com a experiência na loja de produtos personalizados da fintech; ainda que o cliente fosse conectado com a empresa o suficiente para comprar produtos personalizados, a má experiência foi trazida para a rede social de modo negativo. Os textos D e E se referem a supostas falhas no sistema da empresa que poderia gerar uma perda financeira e uma experiência ruim para estes clientes.

Assim, pode-se observar que os usuários se expressam de maneira ampla sobre a relação que possuem com a empresa. Os textos mostram interações desde a usabilidade do aplicativo até a imagem pública que a empresa tem em ações sociais.

- Tweets Positivos Banco

- a) "@stargirlyas Cinemark - meia Bradesco, cliente vivo. São esses esquemas que ainda salvam"
- b) "Feliz Ano Novo @Bradesco (link de uma postagem do banco)"
- c) "Acabou de publicar uma foto em Teatro Bradesco (link da foto)"
- d) "indo escolher um curso na fundação bradesco"
- e) "Sonhei q alguém tinha feito um pix de 7k pra mim, acordei, abri o app do Bradesco, fui ver o saldo e fiquei triste kkkkk"

Já os tweets positivos a respeito do banco não se relacionam tanto à usabilidade dos produtos ou serviços. O texto A mostra a satisfação do cliente por um convênio que o banco tem com uma empresa de cinema. Os textos B e C demonstram uma conexão dos clientes com um banco de maneira mais íntima; isto revela a capacidade do banco de se conectar com estes usuários. O texto D revela o contentamento de um usuário em desfrutar de um conteúdo educativo que o banco disponibiliza na internet. E, por fim, o texto E demonstra, por meio de uma expressão em tom de humor, a preferência pelo banco; isto indica uma boa relação com este cliente, uma vez que o usuário inclui o banco nas interações que gera com os seus pares.

- Tweets Negativos Banco

- a) "Bradesco como q a chave de segurança é inválida se a chave é a minha cara, será se estou feia"
- b) "Uma hora em ligação com o @Bradesco escutando aquela bela música para tentar resolver uma compra não reconhecida, oq o direito do consumidor fala mesmo hein @proconspoficial @BancoCentralBR"
- c) "Novamente, estava tentando pelo Bradesco e não consegui. Tentei pelo @C6Bank e não levou nem 20 segundos. EU AMO OS BANCOS DIGITAIS!"
- d) "JA ENCERREI MINHA CONTA NO BRADESCO (link de uma propaganda de marketing)"
- e) "Pois é Bradesco... agora guenta o tranco e perdas de contas bancárias. (link de uma propaganda de marketing)"

Nos textos negativos, observa-se uma amplitude maior de temas abordados. No texto A, o usuário demonstra em tom jocoso o descontentamento com um dos recursos de segurança do aplicativo. Já o tweet B demonstra a insatisfação de um dos usuários com o atendimento ao consumidor do banco, com indicações de ineficiência no serviço. O texto C mostra que o usuário não conseguiu realizar alguma atividade pelo produto do banco, e também indica a preferência pelas fintechs. O tweet D e E revelam o descontentamento destes usuários com uma determinada propaganda de marketing da empresa durante o período de coleta de



dados. De modo geral, podemos perceber que os tweets negativos são mais abrangentes do que os positivos.

Desta maneira, vemos que as expressões na rede social a respeito de ambas as empresas podem abranger diversos assuntos de modo genérico ou específico. Consegue-se observar grande valor analítico nestes exemplos uma vez que os usuários concedem avaliações sobre aspectos que vão desde a usabilidade até a imagem que têm sobre cada uma das empresas. Muitas vezes essa visão não é refletida em outras vias porque o espaço para descrições detalhadas pode não ser adequado ou também porque os usuários podem não se sentir tão confortáveis como nas redes sociais diante dos seus seguidores.

Para que estes resultados pudessem ter uma relevância maior, optou-se por contextualizá-la com outras vias de análise de satisfação mais consolidadas para que fosse possível entender quais lacunas estes resultados podem preencher.

## **4.2. Análise Exploratória**

A análise exploratória da experiência que os clientes das duas empresas têm com relação ao serviço prestado tem o objetivo de contextualizar e oferecer uma comparação razoável aos resultados obtidos previamente. Portanto, é necessário notar que, segundo as informações que as empresas fornecem, a *fintech* ultrapassou a marca de 41 milhões de clientes no segundo trimestre de 2021 enquanto o banco passou a casa dos 99 milhões de clientes no mesmo período (BACEN, 2021; NUBANK, 2021). Com estes dados, podemos perceber que as duas empresas possuem uma base significativa em relação à população brasileira, com cerca de 210 milhões de pessoas (IBGE, 2021).

Estas informações são imprescindíveis para analisarmos o engajamento e o nível de conexão que estas empresas possuem com os seus clientes por meio das redes sociais, além de nos dar visibilidade dos resultados que ambas têm em vias formalizadas de avaliação que usamos neste estudo. Com isso buscamos ferramentas e plataformas que pudessem ter a representatividade nacional que as empresas conquistaram.

### **4.2.1. Avaliação nas lojas de APPs**

Em sequência, vemos nas figuras 11 e 12 a avaliação que o banco tem nas lojas de APP oficiais. Na figura 11, podemos perceber que o banco teve um total de mais de dois milhões de avaliações e uma concentração expressiva destas em notas 5 (máximo) no Google Play. Na figura 12, na Apple, o banco teve mais de um milhão e meio de avaliações e também uma alta concentração de notas 5 (máximo).



Figura 11 - Avaliações do banco no Google Play (Fonte: Google Play, 2021)



Figura 12 - Avaliações do banco Bradesco na loja da APPs da Apple (Fonte: Loja de Apps Apple, 2021)

E para que a comparação possa ser feita, coletamos as mesmas informações sobre a *fintech*. No mesmo período, as informações das mesmas plataformas foram como exibido nas figuras 13 e 14. Assim, na figura 13, vemos que a empresa digital teve mais de um milhão de avaliações no Google Play e quase 400 mil avaliações na Apple. Em ambas as plataformas, ela obteve uma alta concentração de notas 5 assim como vimos com o banco.

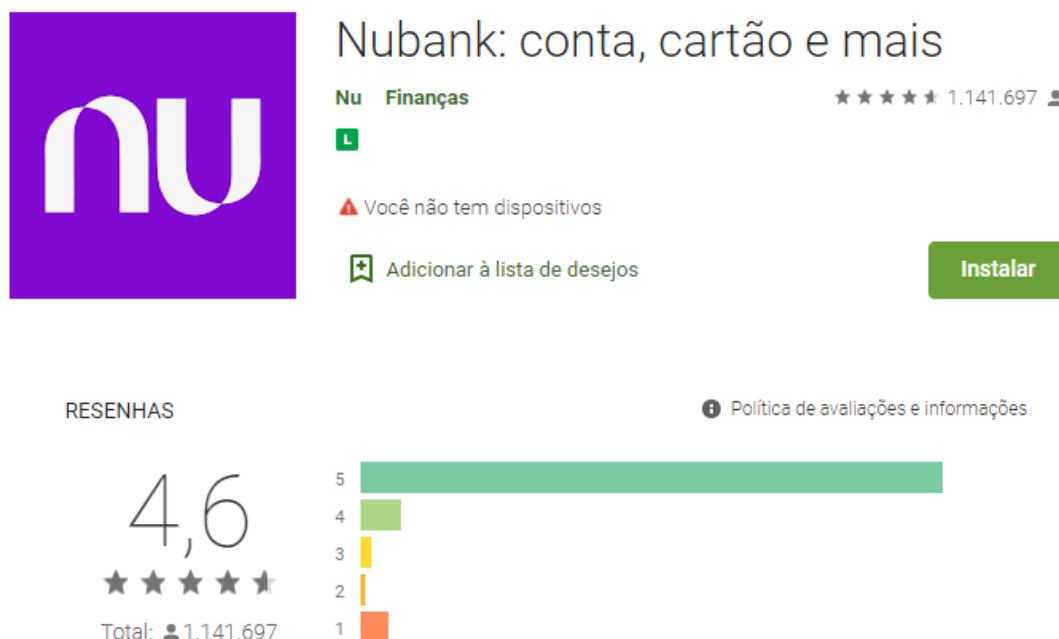


Figura 13 - Figura 13. Avaliações da *fintech* no Google Play Apple (Fonte: Google Play, 2021)



Figura 14 - Avaliações da *fintech* na loja de APPs da Apple (Fonte: Loja de Apps Apple, 2021)

Porém, na média a distribuição das avaliações dá vantagem à *fintech*. Segundo as informações disponibilizadas pelas plataformas, o banco selecionado possui uma soma de notas de 3.815.193 contra uma somatória de 1.533.623 de notas da *fintech*. Desta maneira, podemos considerar que cerca de 3,8% do banco avaliou a experiência do aplicativo em alguma das plataformas enquanto as avaliações da *fintech* chegam a cerca de 3,7% da base de clientes. Assim, podemos concluir que uma proporção semelhante da base de ambas as empresas disponibilizou uma expressão da qualidade do aplicativo na plataforma oficial em que obtiveram o serviço.

Em ambos os meios, na loja Google Play e na loja de APPs da Apple, a *fintech* obtém vantagem em relação ao banco. Na primeira, as notas dão uma média de 4,6 para a *fintech* contra 4,2 para o banco. Na segunda plataforma, a *fintech* fica na frente com 4,8 contra 4,6. Podemos então notar que, mesmo que ambas as empresas tenham notas boas na média das

classificações, os usuários de *fintech* revelam apreciar mais o serviço que recebem do que os clientes do banco. Isto indica que, pelo modelo de negócio totalmente digital, a *fintech* consegue ter uma usabilidade melhor segundo os seus usuários, mas revela também que o banco tradicional conseguiu se adaptar bem a este ambiente.

#### 4.2.2. Google Trends

Com isso, é importante entender a busca que cada empresa tem em um ambiente virtual. A plataforma do Google é usada como busca por assuntos variados e por isso ela consegue revelar o volume que um determinado assunto possui em sua plataforma. Vale ressaltar que, como qualquer tipo de busca conta da mesma maneira, o volume em si não consegue indicar se a busca foi por motivos positivos ou negativos. Com isso obtivemos o resultado como mostra a figura 15. Podemos notar de maneira clara que o banco tem um volume de busca maior do que a empresa digital. Isto está diretamente relacionado com o tamanho da base de clientes, uma vez que o banco tem um volume mais de duas vezes maior. Assim, como o Google tem uma frequência alta de buscas neutras com relação às empresas, este volume se relaciona de maneira direta com o tamanho da base de clientes.

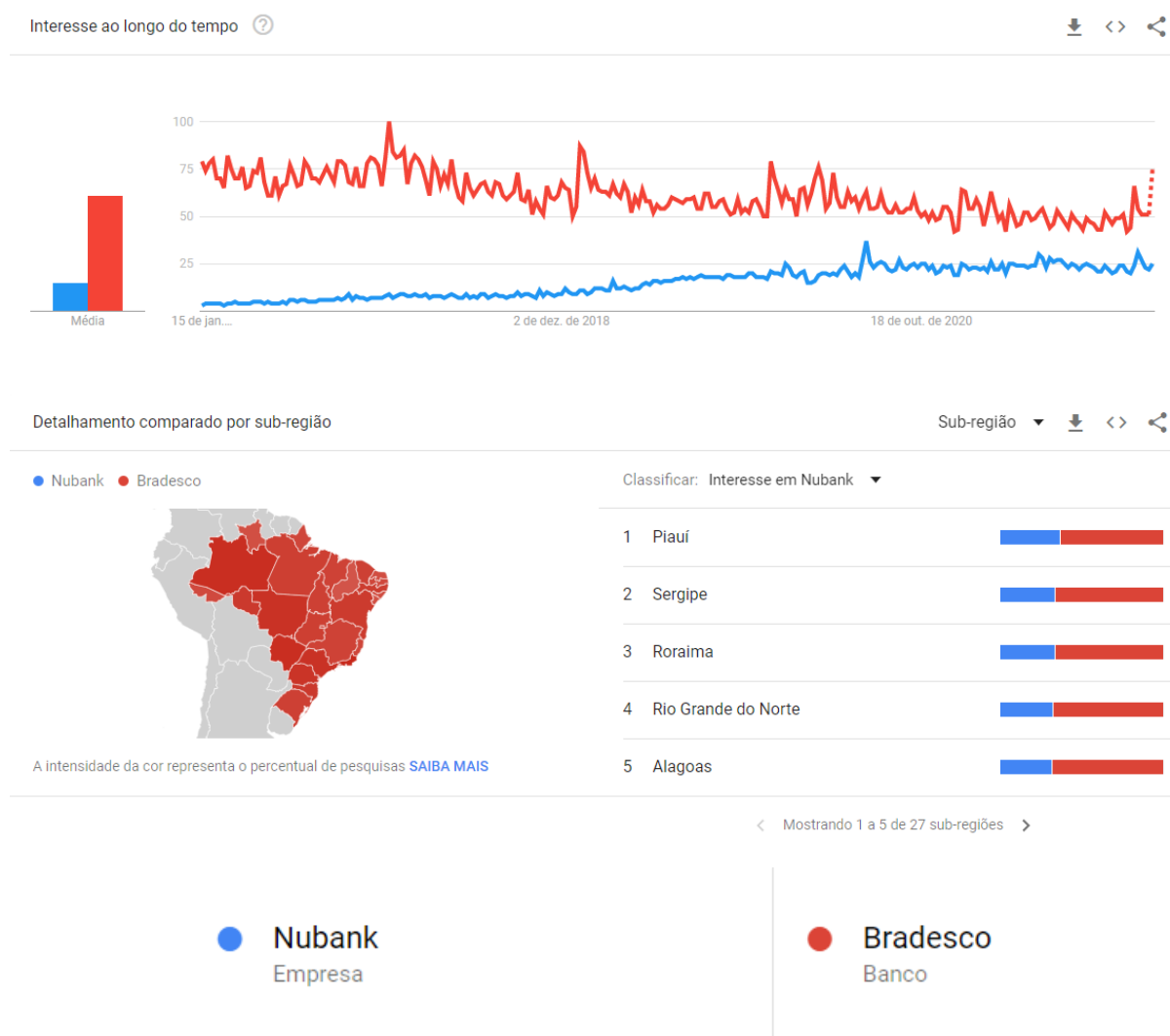


Figura 15 - Buscas no Google Trends (Fonte: Google Trends, 2021)

Porém, a análise de engajamento feita com informações fornecidas pelo Google mostra que o interesse pelo banco tem diminuído em aproximadamente 30% nos últimos cinco anos. No mesmo período, o interesse pela *fintech* selecionada tem crescido em mais de cinco vezes. Com estas informações, podemos ver que existe uma mudança no comportamento de ambas as bases de clientes em relação ao engajamento pelos meios virtuais. Isto ocorre não apenas pelo alto crescimento da *fintech* em número de clientes mas também pelos diversos lançamentos de produtos que o banco teve no mesmo período (NUBANK, 2021). Como o banco tradicional tende a ter um modelo já completo, ele não conta com a atração que novos serviços geram. Isto poderia ser um grande influenciador da grande subida da *fintech*. Ao mesmo tempo, a estabilidade de um modelo de negócio faz com os clientes busquem menos informações ao longo do tempo pela familiaridade que têm com os produtos. Isto pode influenciar diretamente as buscas do banco tradicional.

Vemos também que o banco leva ampla vantagem de busca na comparação por unidades federativas, uma vez que ele tem o volume maior em todas. Porém, isto tende a ser apenas um resumo da situação geral de um maior volume de clientes na base. Desta forma, a comparação tende a dar vantagem ao banco sob a perspectiva da procura, porém uma vantagem para a *fintech* quanto à evolução desta busca.

#### 4.2.3. Reclame Aqui

Para esta análise de tratamento de problemas, utilizamos os dados fornecidos pela empresa Reclame Aqui. Podemos ver na comparação da figura 16 que a *fintech* leva uma ligeira vantagem em relação ao banco. A *fintech*, segundo a empresa, teve um índice de solução 5,5 pontos percentuais maior que o do banco. A empresa digital também obteve uma taxa de 10,2 pontos percentuais maior na avaliação de clientes que voltariam a fazer negócio com a mesma empresa.





Figura 16 – Avaliações de clientes no Reclame Aqui (Fonte: Reclame Aqui, 2021)

Além disso, a empresa digital performou um tempo médio de resposta de aproximadamente um terço do banco. Ao final, a nota média dos consumidores para o atendimento recebido foi 10% mais alta para empresa digital em comparação com o banco. Estes resultados indicam que a *fintech* conseguiu atender em um menor tempo, indicado pelo tempo médio de resposta, e com maior qualidade os clientes que tiveram problemas relevantes dentro do serviço. A empresa digital também teve um melhor desempenho em manter os clientes que, em algum momento, não ficaram contentes com a experiência que tiveram. Em suma, a *fintech* proporcionou uma melhor experiência para os seus clientes ao tratar as situações problemáticas graves, uma vez que todas métricas levantadas pelo Reclame Aqui dão vantagem à empresa digital.

Podemos perceber que a *fintech* possui uma incidência média de reclamações de 0,04% enquanto o banco possui 0,02%. Nesta comparação o banco leva vantagem pois mostra que uma proporção menor da sua base buscou esta plataforma para registrar uma experiência ruim. Porém, como ambos os números são pequenos em comparação com a base, isto pode estar relacionado pelo fato de que, como o banco possui agências, muitos clientes podem ter optado em solucionar o problema presencialmente. Assim, estes resultados revelam mais o potencial que cada empresa tem para trabalhar as situações negativas do que uma avaliação média da experiência de cada serviço.

#### 4.2.4. Pesquisa de NPS

A pesquisa de NPS em que nos baseamos foi feita pela empresa SoluCX e é amplamente aceita no mercado. Nesta pesquisa, como exibido na figura 17, vemos que 64,2% dos entrevistados já tiveram contato com a *fintech* (6ª posição) escolhida enquanto 79,2% das pessoas já tiveram contato com o banco (2ª posição) escolhido. Isto está diretamente relacionado ao volume de clientes que cada empresa possui, uma vez que o banco tinha mais que o dobro de clientes do que a *fintech* durante o período da pesquisa de NPS.

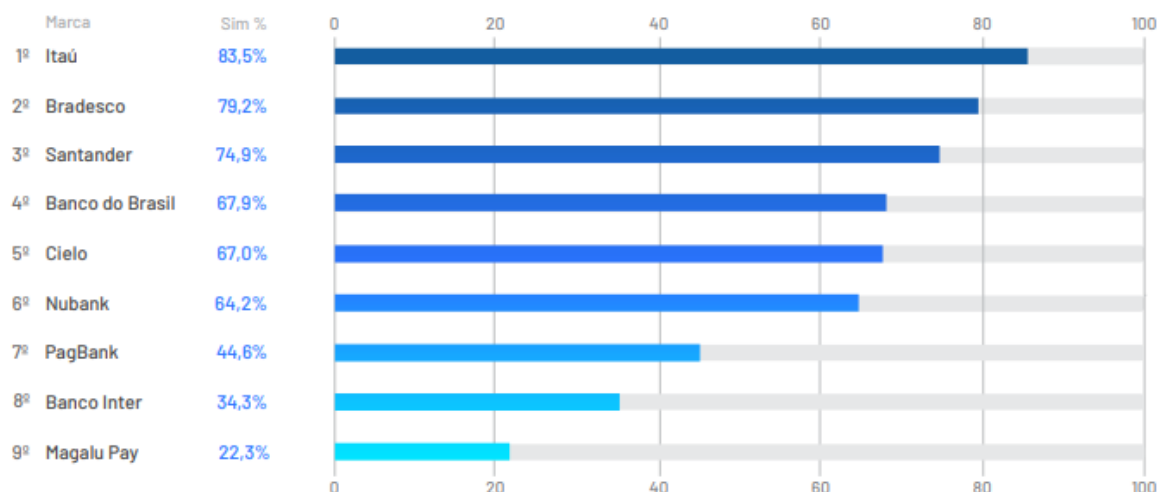


Figura 17 - Ranking de interações dos clientes com empresas financeiras pela SoluCX (Fonte: SoluCX, 2021)

Porém, nesta população entrevistada na pesquisa, a *fintech* teve uma avaliação de 69,4 (1ª posição do ranking) enquanto o banco selecionado teve uma avaliação pelos usuários de apenas 27,0 (7ª posição do ranking) como mostra a figura 18.

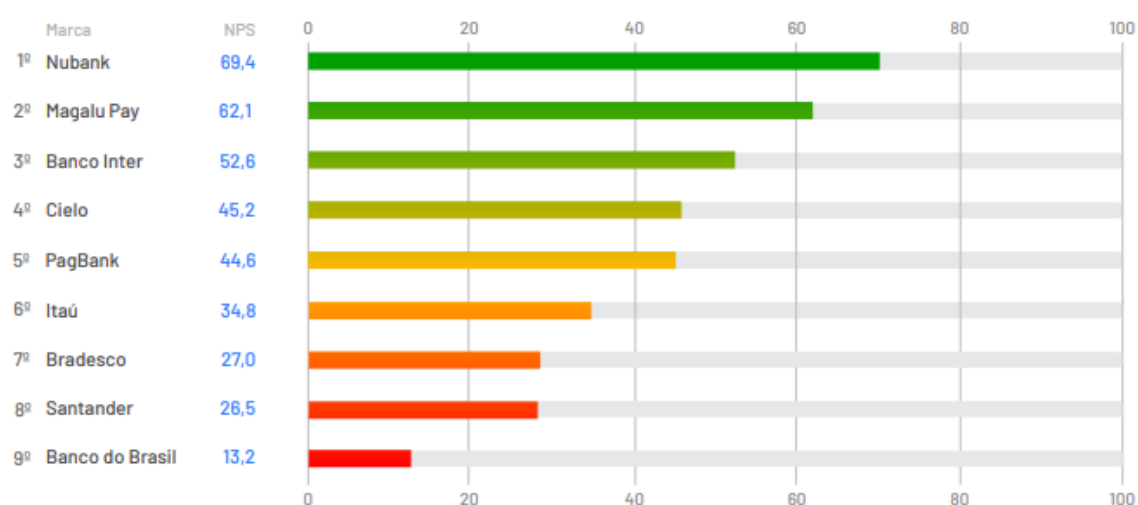


Figura 18 - Ranking de NPS das empresas financeiras (Fonte: SoluCX, 2021)

Desta maneira, conseguimos perceber que, por mais que o banco tenha uma capilaridade maior entre as pessoas, a percepção que se tem da empresa é muito inferior a visão que se tem da *fintech*. Isto revela que, nesta avaliação que leva em conta toda a experiência, relacionamento com a empresa e a percepção da qualidade da marca, a *fintech* teve uma ampla vantagem em mostrar valor aos usuários. Isto indica que a *fintech* proporcionou uma experiência melhor aos seus clientes. Como este índice busca refletir a diferença entre pessoas que vão promover e aquelas que vão depreciar uma determinada empresa, podemos entender que o banco terá uma tendência menor em ser promovida pelos clientes do que a

*fintech*. Isto pode indicar a empresa digital detém uma confiança maior da sua base de clientes, bem como fazer com que ela tenha um crescimento maior no mercado.

Com estes resultados, seguiu-se para entender como ambas as empresas poderiam ser vistas pelos clientes que tiveram problemas durante a experiência com o banco.

#### 4.2.5. Engajamento nas redes sociais Instagram e Twitter

Outra forma que usamos para medir a satisfação e engajamento dos clientes com as empresas foram as redes sociais. Por meio delas, buscamos entender qual das companhias conseguiu estabelecer uma comunicação melhor, bem como fomentar um interesse em acompanhar e interagir com aquilo que elas publicam rotineiramente. Para isso, utilizamos informações de duas redes consolidadas no mercado, Instagram e Twitter, como mostra as figuras 19 e 20:



Figura 19 - Número de clientes do banco Bradesco no Instagram e Twitter (Fonte: Instagram, 2021; Twitter, 2021)

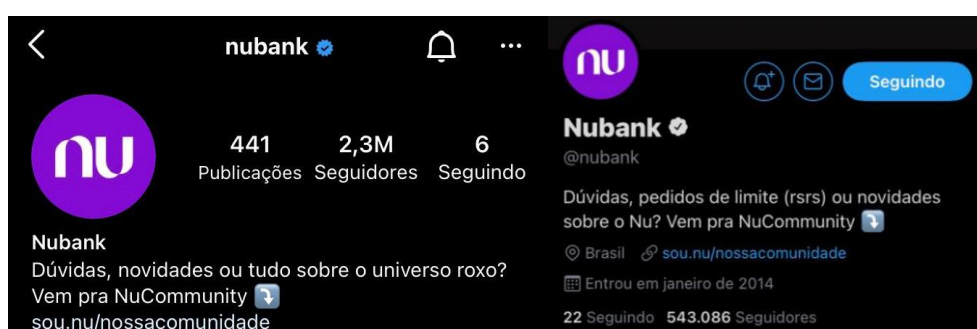


Figura 20 - Número de clientes da *fintech* Nubank no Instagram e no Twitter (Fonte: Instagram, 2021; Twitter, 2021)

As figuras 19 e 20 foram coletadas das plataformas indicadas no dia 30 de dezembro de 2021. Podemos ver que, mesmo o banco em questão tendo cerca de 2,5 vezes o tamanho da *fintech* em número de clientes, o engajamento em duas redes sociais populares é significativamente inferior. No Instagram, a *fintech* conta com 2,3 milhões de seguidores



enquanto o banco conta com apenas com 407 mil; uma relação de mais de 5 vezes o volume de seguidores. No Twitter, a mesma rede da qual coletamos os dados a serem interpretados, a *fintech* conta com 543 mil contra 202 mil do banco; uma relação de mais de 2 vezes.

Desta forma, nota-se que a *fintech* possui um volume maior de seguidores nas principais mídias sociais mesmo com um volume menor de clientes. Isso indica uma maior destreza da *fintech* em se comunicar com os seus clientes pelos meios digitais, bem como uma maior preferência da base de clientes dela em se relacionarem por esses mesmos meios.

#### **4.3. Comparação entre as vias**

Desta mesma forma, o engajamento mostrado pelas redes sociais também indica uma busca mais intensa pelo relacionamento dos clientes com a *fintech* do que com o banco. Isso mostra, portanto, a relevância de se verificar as expressões de satisfação dos clientes por meio destas plataformas e não apenas as vias convencionais que o mercado já utiliza para comparar empresas.

Este fenômeno das redes sociais também pode estar relacionado com uma possível diferença da composição das bases, onde a *fintech* teria uma proporção maior de sua base de clientes que utilizam estes canais do que o banco selecionado. Caso isso seja verificado, aumentaria a relevância de se intensificar o monitoramento das expressões desses usuários nestas redes.

Na análise exploratória, conseguimos perceber que a *fintech* leva vantagem na percepção de seus clientes com relação à experiência fornecida. Primeiro, como vimos na comparação feita pelas avaliações nas lojas de aplicativos, a *fintech* tem notas maiores do que o banco. Isto indica que, pelo caráter da avaliação com foco em quesitos técnicos, a usabilidade do produto é considerada melhor pelos clientes da empresa digital. Em seguida, com a análise feita pela plataforma do Google, notamos que a busca atual pelo banco é maior do que pela *fintech* nesta mesma ferramenta; porém, existe uma nítida queda na busca pela empresa tradicional enquanto existe um aumento ainda mais expressivo na busca pela empresa digitalizada no mesmo período. Uma limitação desta ferramenta é o fato de ela não indicar se a busca é por uma situação favorável ou contrária a empresa.

Em seguida, buscamos analisar informações que contemplassem melhor a experiência que os clientes de ambas as empresas, além dos fatores mais diretos de uso e satisfação. Para isso, analisamos primeiro a capacidade das companhias de lidar com problemas graves por meio da plataforma Reclame Aqui. Nesta plataforma, os clientes podem registrar alguma situação problemática que tiveram e as empresas podem utilizar estas informações para se comunicar com o cliente. Assim, vimos que a *fintech* leva vantagem em termos de atendimento, pois possui um tempo menor de resposta e uma assertividade maior. Desta forma, os próprios clientes indicam a vantagem pela disparidade no índice “voltaria a fazer negócio” no qual a *fintech* possui uma nota maior. De igual forma, a métrica NPS também indica uma vantagem clara da *fintech* em relação ao banco. Na comparação, a nota do banco é menos da metade da *fintech*; 27 para o banco contra 69,4 para a *fintech*. Isso indica que a experiência como um todo tem sido mais agradável na *fintech* do que no banco.

Quando comparamos o engajamento que ambas as empresas possuem vemos que, embora o banco tenha uma busca na plataforma do Google maior, o engajamento é significativamente maior com a *fintech* pelos seus usuários do que pelo banco. A empresa digitalizada possui um número absoluto de seguidores maior em redes sociais como Instagram e Twitter mesmo com uma base de clientes menor do que a metade do banco tradicional. Desta maneira, o

conjunto das informações apresentadas nos revela como a percepção da experiência dos clientes, bem como o relacionamento que eles têm com as empresas, é melhor desenvolvido pela empresa selecionada do ramo digital do que aquela do ramo tradicional bancário. A análise exploratória mostra que a diferença vista nas expressões dos textos do Twitter se concentra mais no relacionamento de confiança e lealdade de usuários com empresas do que nos quesitos de qualidade e atendimento. Porém, a *fintech* leva vantagem em ambas as perspectivas.

## 5. Conclusão

Esta pesquisa teve como objetivo avaliar a experiência de usuários de dois ramos do segmento bancário, o totalmente digital e o tradicional. Este objetivo foi separado em dois complementares. O primeiro era utilizar informações provenientes da rede social Twitter e interpretá-las por um algoritmo de mineração de emoções, Naive Bayes. A segunda parte era utilizar outras vias de avaliação de experiência para tomar como referência os resultados extraídos e também contextualizar as conclusões obtidas. Desta forma, a primeira parte agrega o valor de trazer as expressões mais informais e feitas espontaneamente num ambiente de interações com pessoas próximas. Enquanto isso, a segunda parte traz o valor de avaliações mais formais e consolidadas onde os clientes caracterizam o serviço de maneira estruturada.

Para isso, estruturamos o algoritmo Naive Bayes e coletamos dados do Twitter disponibilizados na própria API da rede. E com isso, obtivemos os resultados de que, no intervalo de tempo selecionado (novembro de 2021 a janeiro de 2022), a *fintech* teve uma frequência de sentimentos positivos maior do que o banco. Realizamos também uma análise de nuvem de palavras, que mostra as palavras com maior frequência na amostra. Com esta análise, conseguimos ver que os clientes se expressam bastante com relação aos produtos oferecidos por ambas as empresas. Os resultados também indicaram que o banco teve críticas negativas com relação à uma propaganda realizada no período da coleta.

Aliada às análises de mineração de textos, exploramos resultados de outras vias de avaliação de experiência e engajamento com as empresas. Estes outros resultados deram ampla vantagem à *fintech* tanto na perspectiva de engajamento como de experiência. Esta análise revelou uma melhor performance da empresa digital em gerar uma comunicação com os clientes nas redes sociais, e uma melhor performance em construir confiança e lealdade. Esta vantagem que a empresa digital obteve nas análises de mineração de textos tende a estar mais concentradas na parte do relacionamento com a marca da empresa do que com requisitos técnicos, ainda que estes também sejam melhor avaliados pelos clientes.

É importante destacar que as duas empresas selecionadas possuem portfólios de produtos em níveis diferentes. O banco possui uma gama maior de produtos complexos de financiamento e linhas de crédito do que a *fintech*. Isto pode ter uma influência significativa na expectativa que os clientes têm da experiência que vão experimentar. Os clientes que têm uma proporção maior das finanças concentradas em uma empresa podem ser mais críticos diante de um mesmo problema. Este viés não foi estudado nesta pesquisa e, portanto, entende-se que pode ser explorado como uma oportunidade de pesquisa.

Estes resultados mostram que a análise de sentimentos pode ter um alto valor a ser agregado com relação ao dinamismo que as redes sociais podem oferecer e que nem sempre as demais vias conseguem acompanhar. Como a avaliação das redes pode ser feito em tempo real, isto pode gerar um valor imensurável de acompanhar o sentimento dos usuários a cada

ação pública e lançamentos de produtos que uma determinada empresa pode realizar. Com esta análise, também é possível levantar o conhecimento de qual subtópico os clientes dão maior foco dentro de um amplo contexto que poderia ser uma empresa, evento ou produto. Desta forma, esta via de avaliação não substitui por completo as demais que já são utilizadas, mas consegue agregar um valor que elas não trazem.

Uma oportunidade de pesquisa que observamos é a criação de métodos para averiguar a confiabilidade da classificação dos textos. A biblioteca Naive Bayes tem opções para julgar a proximidade que a base de comparação tem com a amostra, porém seria necessário verificar também o quão bom é este modelo para classificar textos informais como os estudados, uma vez que o algoritmo pode não ser tão eficiente em avaliar sarcasmos, ironias ou subjetividades presentes em textos informais. Outra oportunidade que vemos para trabalhos futuros é separação de um limiar de sentimentos neutros uma vez que o algoritmo traz uma nota, mas sem um padrão do que é necessariamente neutro ou se encontra em alguma polaridade. Na nossa pesquisa, utilizamos uma referência arbitrária para que pudéssemos realizar os cálculos. Além disso, nota-se como outra limitação de pesquisa as limitações da ferramenta de coleta quanto à quantidade de tweets que podem ser coletados no total e também à distância de tempo em que eles podem ser buscados. Por outro lado, também notamos que a rede não disponibiliza um método oficial da contagem de quantas vezes aquele tema foi referenciado. Com isso, não temos uma maneira eficaz de confirmar a representatividade da base que pegamos para análise. Isto é uma oportunidade que enxergamos para trabalhos futuros nesta área.

## 6. Referências bibliográficas

ABDULLAH, N. A. S.; RUSLI, N. I. A. Multilingual sentiment analysis: A systematic literature review. **Pertanika Journal of Science and Technology**, v. 29, n. 1, p. 445–470, 2021.

ABRAS, A.; MATTOS, G. G. C. Get Them While They Are Young: Social Media and Customer Service at Retail Banks. **Journal of Financial Services Research**, v. 59, n. 1–2, p. 97–113, 2021.

ABRAS, C.; MALONEY-KRICHMAR, D.; PREECE, J. User-Centered Design. p. 1–14, 2004.

ALI, Z. et al. Improving sentiment analysis efficacy through feature synchronization. **Multimedia Tools and Applications**, p. 13325–13338, 2021.

AZARENKOVA, G. et al. The influence of financial technologies on the global financial system stability. **Investment Management and Financial Innovations**, v. 15, n. 4, p. 229–238, 2018.

BANCO CENTRAL DO BRASIL (BACEN). **Relatório de Economia Bancária 2019**. [s.l.: s.n.]. Disponível em: <<https://www.bcb.gov.br/publicacoes/relatorioeconomiabancaria>>.

BANCO CENTRAL DO BRASIL (BACEN). **Informações Gerais - Segmento Bancário**. Disponível em: <<https://www3.bcb.gov.br/ranking/historico.do>>. Acesso em: 30 dez. 2021.

BARCHIESI, M. A.; FRONZETTI COLLADON, A. Big data and big values: When companies need to rethink themselves. **Journal of Business Research**, v. 129, n. September 2018, p. 714–722, 2021.

BIRJALI, M.; KASRI, M.; BENI-HSSANE, A. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. **Knowledge-Based Systems**, v. 226, p. 107134, 2021.

BOLLEN, J.; MAO, H.; ZENG, X. Twitter mood predicts the stock market. **Journal of**

**Computational Science**, v. 2, n. 1, p. 1–8, 2011.

BRADLEY, C. et al. A new perspective on personas and customer journey maps: Proposing systemic UX. **International Journal of Human Computer Studies**, v. 148, n. January, p. 102583, 2021.

CAMPESE, C.; AMARAL, D. C.; MASCARENHAS, J. Restating the Meaning of UCD and HCD for A New World of Design Theories. **Interacting with Computers**, v. 32, n. 1, p. 33–51, 2020.

CHAWLA, D.; JOSHI, H. The Moderating Effect of Demographic Variables on Mobile Banking Adoption: An Empirical Investigation. **Global Business Review**, v. 19, n. 3\_suppl, p. S90–S113, 2018.

CRUZ-GARCÍA, P.; FERNÁNDEZ DE GUEVARA, J.; MAUDOS, J. Bank competition and multimarket contact intensity. **Journal of International Money and Finance**, v. 113, 2021.

DOHERTY, R. A.; SORENSON, P. Keeping Users in the Flow: Mapping System Responsiveness with User Experience. **Procedia Manufacturing**, v. 3, n. Ahfe, p. 4384–4391, 2015.

DUONG, H. T.; NGUYEN-THI, T. A. A review: preprocessing techniques and data augmentation for sentiment analysis. **Computational Social Networks**, v. 8, n. 1, p. 1–16, 2021.

FAHMI, D. Y.; HARTOYO; ZULBAINARNI, N. Mining Social Media (Twitter) Data for Corporate Image Analysis: A Case Study in the Indonesian Mining Industry. **IOP Conference Series: Earth and Environmental Science**, v. 1811, n. 1, 2021.

FERNANDES, E. et al. A data-driven approach to measure restaurant performance by combining online reviews with historical sales data. **International Journal of Hospitality Management**, v. 94, n. March 2020, 2021.

GANDOLFO, A. **Content shared between banks and users on the social ecosystem: an inductive exploratory inquiry**. [s.l.] Springer US, 2020. v. 20

GARZARO, D. M.; VAROTTO, L. F.; PEDRO, S. DE C. Internet and mobile banking: the role of engagement and experience on satisfaction and loyalty. **International Journal of Bank Marketing**, v. 39, n. 1, p. 1–23, 2020.

GOMBER, P.; KOCH, J. A.; SIERING, M. Digital Finance and FinTech: current research and future research directions. **Journal of Business Economics**, v. 87, n. 5, p. 537–580, 2017.

GOZUACIK, N.; SAKAR, C. O.; OZCAN, S. Social media-based opinion retrieval for product analysis using multi-task deep neural networks. **Expert Systems with Applications**, v. 183, n. May, p. 115388, 2021.

GUPTA, A. et al. Comprehensive review of text-mining applications in finance. **Financial Innovation**, v. 6, n. 1, p. 1–25, 2020.

HASSENZAHN, M.; TRACTINSKY, N. User experience - A research agenda. **Behaviour and Information Technology**, v. 25, n. 2, p. 91–97, 2006.

IBGE. **População brasileira chega a 213,3 milhões de habitantes, estima IBGE**. Disponível em: <<https://www.gov.br/pt-br/noticias/financas-impostos-e-gestao-publica/2021/08/populacao-brasileira-chega-a-213-3-milhoes-de-habitantes-estima-ibge>>.

JALILVAND, M. R. et al. Factors influencing word of mouth behaviour in the restaurant industry. **Marketing Intelligence and Planning**, v. 35, n. 1, p. 81–110, 2017.

JAMSHIDI, D. et al. Mobile banking behavior and flow experience: An integration of

utilitarian features, hedonic features and trust. **International Journal of Social Economics**, v. 45, n. 1, p. 57–81, 2018.

JIANG, W. et al. Using geotagged social media data to explore sentiment changes in tourist flow: A spatiotemporal analytical framework. **ISPRS International Journal of Geo-Information**, v. 10, n. 3, 2021.

JOHANN, M.; WOLF, C.; GODULLA, A. Managing relationships on Facebook: A long-term analysis of leading companies in Germany. **Public Relations Review**, v. 47, n. 3, p. 102044, 2021.

KAPOOR, A. P.; VIJ, M. How to Boost your app Store Rating? An Empirical Assessment of Ratings for Mobile Banking Apps. **Journal of Theoretical and Applied Electronic Commerce Research**, v. 15, n. 1, p. 99–115, 2020.

KARAM, E.; HUSSEIN, W.; GHARIB, T. F. Integrating location and textual information for detecting affected people in a crisis. **Social Network Analysis and Mining**, v. 11, n. 1, 2021.

KIM, S. Y. et al. Public sentiment toward solar energy—opinion mining of twitter using a transformer-based language model. **Sustainability (Switzerland)**, v. 13, n. 5, p. 1–19, 2021.

KUJALA, S. Effective user involvement in product development by improving the analysis of user needs. **Behaviour and Information Technology**, v. 27, n. 6, p. 457–473, 2008.

LAW, E. L. C. et al. Understanding, scoping and defining user experience: A survey approach. **Conference on Human Factors in Computing Systems - Proceedings**, n. April, p. 719–728, 2009.

LEE, H. J. et al. Mining service quality feedback from social media: A computational analytics method. **Government Information Quarterly**, v. 38, n. 2, p. 101571, 2021.

MANOJ KUMAR, K. et al. Enhanced text mining methodology in social media platform. **International Journal of Innovative Technology and Exploring Engineering**, v. 8, n. 12, p. 4857–4861, 2019.

MCNAMEE, P.; MENDOLIA, S.; YEROKHIN, O. Social media use and emotional and behavioural outcomes in adolescence: Evidence from British longitudinal data. **Economics and Human Biology**, v. 41, p. 100992, 2021.

MILIAN, E. Z.; SPINOLA, M. DE M.; CARVALHO, M. M. D. Fintechs: A literature review and research agenda. **Electronic Commerce Research and Applications**, v. 34, n. September 2018, 2019.

MUSLIM, E. et al. User interface redesign of e-commerce platform mobile application (Kudo) through user experience evaluation to increase user attraction. **IOP Conference Series: Materials Science and Engineering**, v. 508, n. 1, 2019.

NLTK. **Natural Language Toolkit Documentation**. Disponível em: <<https://www.nltk.org/>>. Acesso em: 31 dez. 2021.

NOYES, D. **Top 10 Twitter Statistics**. Disponível em: <<https://zephoria.com/twitter-statistics-top-ten/>>. Acesso em: 21 jun. 2021.

NUBANK. **Resultado do Nubank Brasil: cada dia mais completo para atender mais de 40 milhões de clientes**. Disponível em: <<https://blog.nubank.com.br/resultado-nubank-2021-primeiro-semester/>>. Acesso em: 30 dez. 2021.

OECD. **PISA 2015 Results (Volume III): Students' Well-Being** Oecd. [s.l: s.n.]. Disponível em: <[https://www.oecd-ilibrary.org/education/pisa-2015-results-volume-iii\\_9789264273856-en](https://www.oecd-ilibrary.org/education/pisa-2015-results-volume-iii_9789264273856-en)>.

- OUERTATANI, A.; GASMI, G.; LATIRI, C. Parsing argued opinion structure in Twitter content. **Journal of Intelligent Information Systems**, v. 56, n. 2, p. 327–353, 2021.
- OWUSU KWATENG, K.; OSEI ATIEMO, K. A.; APPIAH, C. Acceptance and use of mobile banking: an application of UTAUT2. **Journal of Enterprise Information Management**, v. 32, n. 1, p. 118–151, 2019.
- PHAN, H. T. et al. An approach for a decision-making support system based on measuring the user satisfaction level on Twitter. **Information Sciences**, v. 561, p. 243–273, 2021.
- PINDADO, E.; BARRENA, R. Using Twitter to explore consumers' sentiments and their social representations towards new food trends. **British Food Journal**, v. 123, n. 3, p. 1060–1082, 2020.
- PLATANIA, M.; SPADONI, R. How people share information about food: Insights from tweets regarding two Italian Regions. **International Journal on Food System Dynamics**, v. 9, n. 2, p. 149–165, 2018.
- POCHAMPALLY, R.; VARMA, V. User context as a source of topic retrieval in Twitter. ... on **Enriching Information Retrieval** ( ..., n. Enir, p. 1–3, 2011.
- RAY, A.; BALA, P. K.; JAIN, R. Utilizing emotion scores for improving classifier performance for predicting customer's intended ratings from social media posts. **Benchmarking**, v. 28, n. 2, p. 438–464, 2021.
- RAY, A.; BALA, P. K.; RANA, N. P. Exploring the drivers of customers' brand attitudes of online travel agency services: A text-mining based approach. **Journal of Business Research**, v. 128, n. February, p. 391–404, 2021.
- SADAGHEYANI, H. E.; TATARI, F. Investigating the role of social media on mental health. **Mental Health and Social Inclusion**, v. 25, n. 1, p. 41–51, 2020.
- SERRANO-GUERRERO, J.; ROMERO, F. P.; OLIVAS, J. A. Fuzzy logic applied to opinion mining: A review. **Knowledge-Based Systems**, v. 222, p. 107018, 2021.
- SOLUCX. **Benchmarking NPS Bancário**. [s.l.: s.n.].
- STIEGLITZ, S. et al. Social media analytics – Challenges in topic discovery, data collection, and data preparation. **International Journal of Information Management**, v. 39, n. December 2017, p. 156–168, 2018.
- TENG, S.; KHONG, K. W. Examining actual consumer usage of E-wallet: A case study of big data analytics. **Computers in Human Behavior**, v. 121, n. March, p. 106778, 2021.
- THAKOR, A. V. Fintech and banking: What do we know? **Journal of Financial Intermediation**, v. 41, n. August 2019, 2020.
- TWITTER INC. **API Reference**. Disponível em: <<https://docs.tweepy.org/en/v3.5.0/api.html>>. Acesso em: 2 nov. 2021a.
- TWITTER INC. **Getting access to the Twitter API**. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/getting-started/getting-access-to-the-twitter-api>>. Acesso em: 2 nov. 2021b.
- VARDOMATSKYA, L.; KUZNETSOVA, V.; PLOTNIKOV, V. The financial technologies transformation in the digital economy. **E3S Web of Conferences**, v. 244, p. 10046, 2021.
- VERMEEREN, A. P. O. S. et al. User experience evaluation methods: Current state and development needs. **NordiCHI 2010: Extending Boundaries - Proceedings of the 6th Nordic Conference on Human-Computer Interaction**, p. 521–530, 2010.

WEBSTER, J.; WATSON, R. ANALYZING THE PAST TO PREPARE FOR THE FUTURE: WRITING A LITERATURE REVIEW. **Free Radical Biology and Medicine**, v. 38, n. 12, p. 1662–1666, 2005.

WELEKAR, R.; KARANDIKAR, A.; TIRPUDE, S. Emotion categorization using twitter. **International Journal of Advanced Trends in Computer Science and Engineering**, v. 9, n. 3, p. 2698–2701, 2020.

ZHAN, Y. et al. A social media analytic framework for improving operations and service management: A study of the retail pharmacy industry. **Technological Forecasting and Social Change**, v. 163, n. November 2020, p. 120504, 2021.