DA FOTOGRAFIA AO FILME: UMA ABORDAGEM LONGITUDINAL DE DADOS NA AVALIAÇÃO DO POTENCIAL DE CRÉDITO

FROM SNAPSHOT TO TRAJECTORY: A LONGITUDINAL APPROACH TO CREDIT POTENTIAL ASSESSMENT

KARIN THIES
UNIVERSIDADE PRESBITERIANA MACKENZIE

ALEXANDRE NABIL GHOBRIL

GILBERTO PEREZ
UNIVERSIDADE PRESBITERIANA MACKENZIE

Comunicação:
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

DA FOTOGRAFIA AO FILME: UMA ABORDAGEM LONGITUDINAL DE DADOS NA AVALIAÇÃO DO POTENCIAL DE CRÉDITO

Objetivo do estudo
Sistematizar e analisar uma experiência profissional aplicada de utilização da dimensão temporal na engenharia de variáveis e reunir evidências sobre sua pertinência para ampliar a representação da realidade econômico-financeira e a suficiência informacional na avaliação do potencial de crédito.

Relevância/originalidade
O relato amplia a discussão sobre a evolução da avaliação do potencial de crédito ao sistematizar uma alternativa complementar à incorporação de novas fontes e à sofisticação algorítmica: explorar a dimensão temporal dos dados já disponíveis por meio da engenharia de variáveis.

Metodologia/abordagem
Relato técnico qualitativo, exploratório e aplicado, baseado na sistematização de experiência profissional diretamente vivenciada, confrontada com literatura especializada e entrevistas semiestruturadas com profissionais brasileiros de crédito, risco, prevenção a fraudes e ciência de dados, analisadas qualitativamente por meio de análise de conteúdo.

Principais resultados
A experiência evidenciou a aplicabilidade operacional da abordagem; as entrevistas revelaram lacunas informacionais reconhecidas pelos especialistas; e a literatura corroborou a relevância das variáveis, fontes e representação. A convergência sustenta sua pertinência, embora o ganho incremental ainda requeira validação quantitativa.

Contribuições teóricas/metodológicas
O relato sistematiza o tempo como mecanismo natural de enriquecimento semântico das variáveis, distinguindo disponibilidade de históricos de sua exploração longitudinal e propondo transformar relações entre observações sucessivas em atributos capazes de representar propriedades não observáveis em registros isolados.

Contribuições sociais/para a gestão
A abordagem sugere às organizações examinar sistematicamente os históricos já disponíveis antes de recorrer exclusivamente a novas bases ou algoritmos mais complexos, especialmente na avaliação de públicos com renda irregular, informalidade ou baixa densidade de informações convencionais.

Palavras-chave: Avaliação do potencial de crédito, Assimetria informacional, Dados longitudinais, Suficiência informacional, Engenharia de variáveis

FROM SNAPSHOT TO TRAJECTORY: A LONGITUDINAL APPROACH TO CREDIT POTENTIAL ASSESSMENT

Study purpose
To systematize and analyze an applied professional experience using the temporal dimension in feature engineering and gather evidence on its relevance for improving the representation of individuals' economic and financial realities and informational sufficiency in credit potential assessment.

Relevance / originality
The report broadens the discussion on the evolution of credit potential assessment by systematizing an alternative complementary to new data sources and algorithmic sophistication: exploring the temporal dimension of already available data through feature engineering.

Methodology / approach
A qualitative, exploratory, applied technical report based on the systematization of directly experienced professional practice, confronted with specialized literature and semi-structured interviews with Brazilian professionals in credit, risk, fraud prevention and data science, qualitatively examined through content analysis.

Main results
Professional experience demonstrated operational applicability; interviews revealed informational gaps recognized by practitioners; and literature corroborated the relevance of variables, data sources and representation. Their convergence supports the approach's relevance, although its incremental contribution still requires quantitative validation.

Theoretical / methodological contributions
The report systematizes time as a natural mechanism for semantic enrichment of variables, distinguishing historical-data availability from longitudinal exploration and proposing the transformation of relationships among successive observations into attributes representing properties not observable in isolated records.

Social / management contributions
The approach suggests that organizations systematically examine already available historical data before relying exclusively on new data sources or more complex algorithms, particularly when assessing individuals with irregular income, informal economic activity or limited conventional financial information.

Keywords: Credit potential assessment, Information asymmetry, Longitudinal data, Informational sufficiency, Feature engineering

DA FOTOGRAFIA AO FILME: UMA ABORDAGEM LONGITUDINAL DE DADOS NA AVALIAÇÃO DO POTENCIAL DE CRÉDITO

1 INTRODUÇÃO

Para que o mercado de crédito se expanda de forma sustentável, os seus agentes precisam distinguir, com precisão crescente, indivíduos com maior potencial de adimplência futura daqueles com maior probabilidade de inadimplência. Em economias marcadas por elevada informalidade, heterogeneidade de perfis ocupacionais e limitações na mensuração da renda, essa tarefa permanece desafiadora, tornando a disponibilidade de informações relevantes um dos principais determinantes da qualidade da decisão de crédito. Sob essa perspectiva, a assimetria informacional permanece como um dos principais obstáculos ao funcionamento eficiente dos mercados financeiros, contribuindo para seleção adversa, racionamento de crédito e precificação inadequada do risco (Stiglitz & Weiss, 1981).

Desde meados do século XX, o mercado brasileiro de crédito passou por sucessivas transformações institucionais, regulatórias e tecnológicas que ampliaram progressivamente as informações disponíveis para a avaliação dos tomadores. A implantação do Sistema de Informações de Crédito (SCR), a expansão do Cadastro Positivo, a consolidação do *Open Finance* e a infraestrutura de pagamentos instantâneos baseada no Pix constituem exemplos desse movimento, associado aos esforços para reduzir assimetrias informacionais e aperfeiçoar os processos de avaliação (Banco Central do Brasil, 2024c; 2024b; 2024e; Febraban, 2024). Paralelamente, o desenvolvimento de técnicas estatísticas, computacionais e de aprendizado de máquina ampliou significativamente a capacidade analítica disponível às instituições financeiras.

Apesar desses avanços, parcela relevante das informações utilizadas na avaliação do potencial de crédito continua representando predominantemente a condição econômico-financeira do candidato em determinado momento. Renda atual, ocupação, endividamento e momento financeiro constituem informações relevantes, mas sua observação isolada oferece essencialmente uma fotografia. Aspectos como continuidade da geração de renda, estabilidade ocupacional, evolução profissional, regularidade das entradas financeiras e capacidade de recuperação diante de períodos adversos possuem natureza dinâmica e podem permanecer insuficientemente representados por avaliações predominantemente transversais.

É nesse espaço que se situa a oportunidade explorada neste relato técnico. Em vez de considerar apenas a incorporação de novas fontes de dados ou o aprimoramento das técnicas utilizadas para processá-los, propõe-se investigar em que medida a dimensão temporal, incorporada à engenharia de variáveis, pode contribuir para representar aspectos da trajetória econômico-financeira dos indivíduos que uma observação pontual não permite apreender adequadamente. A metáfora “da fotografia ao filme” sintetiza essa mudança de perspectiva: não se trata de substituir as informações tradicionalmente empregadas, mas de examinar o que pode ser acrescentado à avaliação quando a condição presente é interpretada à luz de sua trajetória.

A oportunidade investigada teve origem em experiência profissional aplicada e foi posteriormente sistematizada e submetida à confrontação com outras fontes de evidência. Para tanto, o relato articula revisão da literatura especializada e entrevistas semiestruturadas com profissionais experientes nas áreas de crédito, risco, prevenção a fraudes e ciência de dados. Os depoimentos foram submetidos à análise de conteúdo e confrontados com o referencial teórico, buscando verificar se as limitações identificadas na prática encontravam correspondência na literatura e na percepção de outros profissionais do mercado.

Diante desse contexto, o objetivo deste relato técnico é sistematizar e analisar uma experiência profissional aplicada relacionada à utilização da dimensão temporal na engenharia de variáveis e reunir evidências teóricas e empíricas sobre sua pertinência para ampliar a suficiência informacional nos processos de avaliação do potencial de crédito. Não se pretende, neste trabalho, isolar quantitativamente o ganho incremental de variáveis longitudinais, mas examinar a consistência conceitual, a relevância prática e o potencial de desenvolvimento dessa abordagem.

Além desta introdução, o relato está estruturado em quatro seções. A seção 2 apresenta o referencial teórico, abordando a assimetria informacional, a evolução das fontes de informação no mercado brasileiro, a relação entre evolução dos algoritmos e evolução da informação e, finalmente, a perspectiva longitudinal da realidade econômico-financeira. A seção 3 descreve os procedimentos metodológicos adotados. A seção 4 caracteriza a experiência profissional relatada e analisa as evidências obtidas, confrontando-as com a literatura. Por fim, a seção 5 apresenta as considerações finais, contribuições, limitações e desdobramentos técnico-profissionais.

2 REFERENCIAL TEÓRICO

2.1 Assimetria informacional e avaliação de crédito

A concessão de crédito constitui essencialmente um problema de decisão sob incerteza. Antes da contratação de uma operação, o credor precisa estimar a capacidade futura de pagamento do tomador sem poder observá-la diretamente, dispondo apenas de informações que representam, de maneira necessariamente parcial, sua realidade econômica, financeira e comportamental. A qualidade da decisão depende, portanto, da capacidade dessas informações de reduzir a incerteza acerca do comportamento futuro, tornando a suficiência informacional um elemento central nos processos de avaliação do potencial de crédito.

Sob a perspectiva da teoria econômica, essa limitação está associada à assimetria informacional, uma das principais imperfeições dos mercados de crédito. No modelo clássico desenvolvido por Stiglitz e Weiss (1981), credores e candidatos dispõem de níveis distintos de informação acerca das condições relacionadas ao cumprimento futuro das obrigações. Enquanto os próprios indivíduos conhecem melhor sua situação econômica, seus projetos e sua disposição para honrar compromissos financeiros, as instituições precisam inferir essas características a partir de evidências observáveis e imperfeitas. Essa diferença de informação afeta tanto a oferta quanto as condições de acesso ao crédito.

Segundo Stiglitz e Weiss (1981), a impossibilidade de distinguir perfeitamente indivíduos de maior e menor risco pode levar os credores a adotar políticas mais conservadoras de concessão. Nessas circunstâncias, aumentos nas taxas de juros não necessariamente compensam o risco adicional assumido, uma vez que podem atrair proporcionalmente tomadores de maior risco – fenômeno associado à seleção adversa – e alterar o comportamento dos próprios tomadores após a contratação. Como consequência, as instituições podem restringir a oferta mesmo diante da existência de demanda, caracterizando o racionamento de crédito. Assim, indivíduos potencialmente capazes de honrar suas obrigações podem permanecer excluídos ou receber condições menos favoráveis não necessariamente por ausência de capacidade de pagamento, mas pela insuficiência de informações capazes de evidenciá-la adequadamente.

Essa discussão permanece relevante mesmo diante da intensa evolução das técnicas analíticas utilizadas pelas instituições financeiras. Conforme destacam Thomas, Edelman e Crook (2017), os modelos de avaliação de crédito constituem instrumentos destinados a reduzir

a incerteza inerente às decisões de concessão mediante a utilização sistemática de informações capazes de representar características relevantes dos tomadores. Técnicas estatísticas, modelos computacionais e algoritmos de aprendizado de máquina não observam diretamente o risco ou a capacidade futura de pagamento; operam sobre variáveis que representam, com diferentes graus de precisão, os fenômenos de interesse.

A questão informacional, portanto, não se restringe à disponibilidade de dados. Um conjunto extenso de informações pode ainda oferecer uma representação incompleta do indivíduo quando os atributos disponíveis não capturam adequadamente características relevantes para a decisão. Neste relato, essa distância entre aquilo que seria relevante conhecer sobre o tomador e aquilo que as informações disponíveis efetivamente permitem representar é tratada como insuficiência informacional. Quanto maior essa distância, maior permanece a parcela de incerteza que precisa ser absorvida pelo processo decisório, independentemente da sofisticação das técnicas empregadas para processar os dados.

A insuficiência informacional pode produzir erros em ambas as direções. Indivíduos com capacidade consistente de pagamento podem ter seu risco superestimado quando as informações disponíveis não permitem diferenciá-los adequadamente, restringindo seu acesso ao crédito ou tornando-o excessivamente oneroso. Em sentido inverso, candidatos de maior risco podem ser avaliados de maneira excessivamente favorável quando fatores relevantes não são suficientemente representados. Reduzir essa lacuna interessa, portanto, simultaneamente às instituições, que buscam maior capacidade de discriminação, e aos tomadores, especialmente aqueles cuja realidade econômica é menos adequadamente representada pelos mecanismos convencionais de avaliação.

A ampliação da suficiência informacional exige, assim, mais do que acumular dados: requer melhorar a representação dos fenômenos relevantes à decisão. Isso pode ocorrer pela incorporação de informações anteriormente indisponíveis, mas também pela construção de formas mais adequadas de representar a realidade a partir das informações acessíveis. A questão passa a ser, portanto, como ampliar aquilo que os processos de avaliação conseguem conhecer e representar sobre o tomador. A evolução das fontes de informação utilizadas pelo mercado brasileiro de crédito, examinada na seção seguinte, constitui uma primeira dimensão dessa busca.

2.2 Evolução das fontes de informação na análise de crédito no Brasil

Se a qualidade da avaliação do potencial de crédito depende da suficiência informacional disponível, torna-se relevante compreender como o mercado buscou ampliá-la ao longo do tempo. No contexto brasileiro, a evolução dos processos de avaliação pode ser interpretada como uma trajetória de incorporação sucessiva de mecanismos institucionais, regulatórios e tecnológicos destinados a reduzir a incerteza e tornar observáveis aspectos anteriormente pouco acessíveis da realidade econômico-financeira dos indivíduos. Do conhecimento construído no relacionamento direto entre credor e tomador aos atuais ecossistemas digitais, cada etapa ampliou, por mecanismos distintos, o conjunto de informações passíveis de utilização nos processos decisórios (Thomas et al., 2017).

Historicamente, os primeiros mecanismos de avaliação apoiavam-se predominantemente no relacionamento direto entre instituições financeiras e clientes, privilegiando conhecimento pessoal, referências comerciais e histórico de relacionamento. Informações acumuladas no próprio vínculo com o cliente constituíam importante recurso para reduzir a incerteza, mas permaneciam essencialmente restritas à instituição que mantinha esse relacionamento. Com a expansão do crédito massificado e a necessidade de padronizar decisões em maior escala, desenvolveram-se sistemas estruturados de compartilhamento de informações cadastrais e registros de inadimplência, posteriormente consolidados pelos *bureaux* de crédito.

Essa transformação permitiu às instituições observar aspectos da experiência do candidato para além de seu relacionamento direto com o credor, ampliando a base informacional disponível para a decisão (Sprague, 1946; Lawrence, 1987; Thomas, Edelman, & Crook, 2017)

Nas décadas seguintes, iniciativas institucionais e regulatórias ampliaram progressivamente essa visibilidade. A criação do Sistema de Informações de Crédito (SCR), administrado pelo Banco Central do Brasil (BaCen), proporcionou às instituições financeiras acesso estruturado a informações consolidadas sobre operações de crédito dos clientes, ampliando a observação de seu endividamento e exposição ao sistema financeiro (Banco Central do Brasil, 2024c). Posteriormente, o Cadastro Positivo incorporou informações relacionadas ao comportamento de pagamento, permitindo que registros de adimplência complementassem a visão tradicionalmente concentrada em ocorrências de inadimplência (Souto, 2019; Banco Central do Brasil, 2024b).

Mais recentemente, o *Open Finance* ampliou substancialmente a diversidade de informações potencialmente acessíveis aos processos de avaliação, possibilitando, mediante consentimento do cliente, o compartilhamento padronizado de dados bancários, financeiros e transacionais anteriormente restritos às instituições de origem (Banco Central do Brasil, 2024c). Além de ampliar o universo informacional acessível, essa infraestrutura introduziu uma mudança relevante na forma de obtenção dos dados ao atribuir ao próprio consumidor maior protagonismo sobre seu compartilhamento. Como as autorizações são concedidas para relacionamentos financeiros específicos, o cliente pode decidir quais instituições integrarão o conjunto de informações disponibilizado ao novo participante. Diferentes escolhas podem, portanto, tornar visíveis diferentes parcelas de sua vida financeira, em consonância com o princípio de "colocar o consumidor no controle" associado ao sistema financeiro aberto (Open Finance Brasil, 2022).

Esse protagonismo acrescenta uma particularidade ao problema informacional enfrentado pelas instituições. Um consumidor que mantenha relacionamentos com diferentes instituições pode compartilhar aquele que concentre sua movimentação cotidiana, sua conta-salário, seus investimentos ou determinado financiamento, mas também pode optar por não compartilhar outros relacionamentos. Consequentemente, o conjunto recebido por uma instituição não corresponde necessariamente à totalidade da vida financeira do candidato. Para a avaliação do potencial de crédito, passa a ser relevante não apenas analisar as informações compartilhadas, mas compreender sua abrangência e representatividade dentro do conjunto de relacionamentos financeiros mantidos pelo indivíduo.

A consolidação da infraestrutura de pagamentos instantâneos baseada no Pix acrescentou outro componente a esse ecossistema informacional. Embora concebido primordialmente como infraestrutura de pagamentos, o ambiente associado ao Pix e ao Diretório de Identificadores de Contas Transacionais (DICT) ampliou as possibilidades de construção de atributos relacionados à movimentação financeira, aos relacionamentos bancários e a determinados aspectos operacionais, oferecendo novas informações potencialmente complementares às utilizadas nos processos de avaliação (Banco Central do Brasil, 2024d).

Paralelamente, a literatura recente destaca a crescente utilização de dados alternativos (*alternative data*) como complemento às fontes tradicionais. Informações provenientes de plataformas digitais, pagamentos eletrônicos, comportamento transacional e outras bases não convencionais apresentam potencial para ampliar a capacidade preditiva dos modelos, sobretudo em segmentos caracterizados por histórico de crédito limitado ou elevada informalidade econômica (Jagtiani & Lemieux, 2019; Hurley & Adebayo, 2016; Frost, Gambacorta, Huang, Shin, & Zbinden, 2019; Hlongwane, Ramaboa, & Mongwe, 2024). Em vez de substituir os mecanismos tradicionais, essas fontes ampliam a possibilidade de observar

aspectos da realidade de indivíduos insuficientemente representados pelas informações convencionais.

Observada em conjunto, essa trajetória revela um movimento recorrente de expansão daquilo que as instituições conseguem observar sobre os candidatos a crédito. O relacionamento direto permitia conhecer aquilo que se manifestava na relação entre cliente e credor; os *bureaux* ampliaram essa visão mediante o compartilhamento entre agentes privados; o SCR tornou acessível uma visão consolidada das exposições ao crédito; o Cadastro Positivo acrescentou informações de adimplência; o Open Finance permitiu ao consumidor compartilhar dados mantidos em outras instituições; Pix e DICT ampliaram as possibilidades associadas aos ecossistemas transacionais; e os dados alternativos incorporaram dimensões provenientes de fontes não convencionais. A Tabela 1 sintetiza essa trajetória.

Tabela 1 – Evolução das fontes de informação na análise de crédito sob a perspectiva da ampliação da suficiência informacional

Paradigma de acesso à informação	Etapa da evolução	Contribuição para a suficiência informacional
Relacionamento direto entre credor e tomador	Crédito relacional	Permite conhecer reputação e histórico local do tomador
Compartilhamento estruturado entre agentes privados	*Bureaux* de crédito	Padroniza e amplia a avaliação para além do relacionamento direto
Compartilhamento institucional promovido pelo regulador	SCR (BaCen)	Amplia a visão sobre endividamento e exposição consolidada ao crédito
Compartilhamento colaborativo do comportamento de pagamento	Cadastro Positivo	Complementa a visão baseada apenas em inadimplência
Compartilhamento consentido pelo próprio consumidor	*Open Finance*	Amplia o acesso a dados bancários, financeiros e transacionais mantidos em outras instituições
Aproveitamento de ecossistemas transacionais	Pix e DICT (Diretório de Identificadores de Contas Transacionais)	Possibilita a construção de atributos associados a relacionamentos e dinâmica transacional
Aproveitamento de fontes não convencionais	Dados alternativos	Complementam informações tradicionalmente indisponíveis para públicos pouco representados

Fonte: Elaborada pelos autores com base em Sprague (1946), Lawrence (1987), Thomas et al. (2017), Souto (2019), Banco Central do Brasil (2024b, 2024c, 2024d, 2024e), Open Finance Brasil (2022), Jagtiani e Lemieux (2019), Hurley e Adebayo (2016) e Hlongwane et al. (2024).

A evolução sintetizada na Tabela 1 mostra que a resposta histórica à insuficiência informacional esteve fortemente associada à ampliação do campo de observação: compartilhar informações antes restritas a uma instituição, consolidar exposições dispersas pelo sistema, incorporar comportamentos positivos de pagamento, permitir acesso consentido a outros relacionamentos financeiros e explorar fontes até então não convencionais. Cada avanço tornou observável algo que anteriormente estava indisponível, disperso ou restrito a determinado agente.

Essa trajetória, contudo, suscita uma questão adicional. Se sucessivas inovações ampliaram aquilo que se consegue observar, até que ponto novos avanços na avaliação do potencial de crédito dependem necessariamente de observar mais? Antes de examinar essa questão sob a perspectiva temporal, é necessário considerar outra dimensão da evolução recente da atividade: em que medida o aprimoramento da avaliação depende da crescente sofisticação

dos algoritmos e em que medida depende da própria qualidade da informação e de sua representação. Esse é o objeto da seção seguinte.

2.3 Da evolução dos algoritmos à evolução da informação

A evolução da avaliação de crédito nas últimas décadas foi acompanhada pelo desenvolvimento de técnicas estatísticas, computacionais e de aprendizado de máquina progressivamente mais sofisticadas. Métodos tradicionalmente empregados, como regressão logística, análise discriminante e árvores de decisão, passaram a conviver com algoritmos capazes de explorar relações não lineares e grandes volumes de dados, incluindo *random forest*, *gradient boosting* e redes neurais artificiais. Esses avanços ampliaram significativamente a capacidade analítica disponível às instituições financeiras e contribuíram para o aprimoramento dos processos de avaliação de risco (Thomas et al., 2017).

Entretanto, a literatura especializada indica que o desempenho dos modelos não depende exclusivamente da sofisticação dos algoritmos empregados. Em uma das revisões comparativas mais abrangentes da área, Lessmann et al. (2015) analisaram diferentes técnicas utilizadas na estimação do risco de crédito e demonstraram que a escolha do algoritmo constitui apenas uma das dimensões associadas ao desempenho preditivo. De forma complementar, a revisão sistemática de Louzada et al. (2016), evidencia a relevância da qualidade dos dados, da seleção e construção de atributos e da capacidade das variáveis de representar adequadamente os fenômenos que se pretende modelar.

Essa interpretação encontra respaldo em Hand e Henley (1997), para quem a capacidade discriminatória dos modelos de *credit scoring* está estreitamente relacionada à qualidade e à adequação das variáveis utilizadas para representar as características dos indivíduos avaliados. Modelos analíticos não observam diretamente os fenômenos de interesse; processam atributos construídos para representá-los. Assim, mesmo técnicas estatísticas, computacionais ou de aprendizado de máquina altamente sofisticadas permanecem condicionadas ao conteúdo informacional e à capacidade representativa das variáveis que recebem como entrada.

Sob essa perspectiva, o aprimoramento dos modelos pode ocorrer por caminhos complementares. O primeiro consiste no desenvolvimento de algoritmos capazes de identificar relações e padrões mais complexos nos dados. O segundo envolve a engenharia de variáveis, pela qual dados disponíveis são selecionados, combinados ou transformados em atributos mais aderentes aos fenômenos de interesse. O terceiro decorre da incorporação de novas fontes de informação, capazes de tornar observáveis aspectos anteriormente ausentes dos processos de avaliação. Em conjunto, essas dimensões determinam não apenas quanto o modelo consegue processar, mas também o que efetivamente lhe é possível conhecer e representar sobre o candidato.

Thomas et al. (2017) reforçam essa interpretação ao relacionar a evolução da avaliação de crédito tanto aos avanços metodológicos quanto à progressiva ampliação das informações disponíveis. Como discutido na seção anterior, informações cadastrais, registros de inadimplência, históricos de pagamento, exposição ao crédito, dados transacionais e fontes alternativas ampliaram sucessivamente o campo de observação dos agentes de crédito. A incorporação dessas fontes, entretanto, somente se converte em capacidade analítica quando as informações são representadas por atributos adequados aos fenômenos que se pretende estimar.

A engenharia de variáveis ocupa, nesse processo, uma posição particularmente relevante. Enquanto os algoritmos determinam como relações presentes nos atributos serão identificadas e combinadas, a engenharia de variáveis participa da definição daquilo que será apresentado ao modelo como representação da realidade. Uma mesma base de dados pode, portanto, oferecer diferentes conteúdos informacionais conforme as variáveis construídas a

partir dela. A evolução da capacidade de avaliação não depende apenas de processar melhor os dados, mas também de representá-los melhor.

A literatura permite, assim, compreender a evolução dos modelos de avaliação de crédito a partir de três dimensões complementares: sofisticação dos algoritmos, engenharia de variáveis e incorporação de novas fontes de informação. Essas dimensões não competem entre si; interagem na determinação da qualidade dos modelos e da suficiência informacional disponível aos processos decisórios. A Tabela 2 sintetiza essa interpretação.

Tabela 2 – Dimensões da evolução dos modelos de avaliação de crédito segundo a literatura

Mecanismo de aprimoramento	Dimensão de evolução	Exemplos
Ampliação da capacidade de identificar relações e padrões nos dados	Algoritmos	Regressão logística, árvores de decisão, *random forest* , *gradient boosting* , redes neurais
Transformação de dados em atributos mais representativos dos fenômenos de interesse	Engenharia de variáveis	Construção e combinação de atributos derivados
Ampliação das dimensões observáveis da realidade econômico-financeira	Novas fontes de informação	Open Finance, dados alternativos, dados públicos

Fonte: Elaborada pelos autores com base em Hand e Henley (1997), Lessmann et al (2015), Louzada et al. (2016), Thomas et al. (2017), Hurley e Adebayo (2016), Jagtiani e Lemieux (2019) e Hlongwane et al. (2024).

A síntese apresentada na Tabela 2 permite distinguir duas respostas recorrentes aos limites dos processos de avaliação: ampliar aquilo que pode ser observado, mediante novas fontes de informação, e aprimorar a maneira de processar e representar aquilo que já está disponível, mediante algoritmos e engenharia de variáveis. A literatura sugere, portanto, que avanços na avaliação do potencial de crédito não dependem necessariamente da substituição contínua de técnicas por algoritmos progressivamente mais sofisticados.

Essa constatação abre uma possibilidade adicional. Se a engenharia de variáveis é capaz de produzir representações diferentes a partir dos mesmos dados, seria possível ampliar a suficiência informacional sem necessariamente incorporar uma nova fonte, explorando relações ainda não representadas nas informações já disponíveis? Mais especificamente, quando uma base contém observações históricas de um mesmo indivíduo, haveria informação potencial nas relações entre essas observações que não está presente em cada registro considerado isoladamente? É essa questão que conduz à perspectiva longitudinal examinada na seção seguinte.

2.4 A perspectiva longitudinal da realidade econômico-financeira

As seções anteriores mostraram dois movimentos complementares na evolução da avaliação do potencial de crédito. De um lado, novas fontes de informação ampliaram progressivamente aquilo que as instituições conseguem observar sobre os candidatos. De outro, o desenvolvimento de técnicas estatísticas, computacionais e de aprendizado de máquina aumentou a capacidade de identificar relações e padrões nos dados disponíveis. A literatura indica, entretanto, que o desempenho dos modelos permanece condicionado à capacidade das variáveis de representar adequadamente os fenômenos de interesse. É nesse ponto que a dimensão temporal introduz uma possibilidade adicional.

Grande parte das informações utilizadas nos processos de avaliação descreve a condição econômico-financeira do candidato em determinado instante. Renda atual, ocupação, endividamento, movimentação financeira ou relacionamento bancário podem apresentar

elevado conteúdo informacional, mas, quando observados pontualmente, representam essencialmente uma fotografia: mostram como o indivíduo se apresenta no momento da avaliação, sem necessariamente revelar o processo que o conduziu até aquela condição.

Duas pessoas podem, por exemplo, apresentar renda semelhante, ausência de restrições cadastrais e padrões equivalentes de movimentação financeira no momento da solicitação de crédito e, ainda assim, ter chegado a essa fotografia por caminhos substancialmente diferentes. Uma pode apresentar uma condição que se manteve de forma consistente ao longo do tempo; outra, a mesma condição após sucessivas oscilações ou uma mudança recente. A fotografia é semelhante. O que as diferencia está no filme – na sequência de acontecimentos que produziu o estado observado.

Essa distinção conduz a um ponto fundamental: dispor de dados históricos não significa, por si só, adotar uma perspectiva longitudinal. Uma base pode conter dezenas ou centenas de observações sucessivas de um mesmo indivíduo e ainda assim utilizá-las apenas como registros independentes ou fotografias acumuladas. A perspectiva longitudinal surge quando as relações entre essas observações passam a ser exploradas deliberadamente pela engenharia de variáveis, permitindo representar propriedades que não são diretamente observáveis em nenhum registro isolado.

É justamente nessa passagem das fotografias acumuladas para o filme que o tempo adquire significado analítico. Quando incorporado à engenharia de variáveis, ele deixa de funcionar apenas como referência cronológica e passa a atuar como um mecanismo natural de enriquecimento semântico das informações. Observações isoladas descrevem estados; suas relações ao longo do tempo podem revelar continuidade, tendência, estabilidade, recorrência, volatilidade, ruptura, recuperação ou adaptação. Essas propriedades não pertencem propriamente a uma observação individual: emergem da relação entre elas.

Sob essa perspectiva, o tempo também pode ser compreendido como um mecanismo de destilação da informação. Uma sucessão potencialmente extensa de registros elementares pode ser condensada, por meio da engenharia de variáveis, em atributos sintéticos dotados de significado econômico. Uma sequência de valores de renda, por exemplo, deixa de representar apenas uma coleção de valores mensais quando dela são extraídas medidas de tendência, estabilidade, volatilidade, períodos de interrupção ou recuperação. Os dados originais permanecem os mesmos; o que se modifica é aquilo que se consegue conhecer a partir deles. É nesse sentido que o tempo pode transformar dados históricos em informação nova.

A Tabela 3 sintetiza essa passagem da fotografia para o filme, contrastando aquilo que pode ser observado em uma avaliação transversal com propriedades que podem tornar-se representáveis quando a dimensão temporal é incorporada à engenharia de variáveis.

Tabela 3 – Da fotografia ao filme: como a perspectiva longitudinal amplia a representação da realidade econômico-financeira

Situação analisada	Abordagem transversal ("fotografia")	Perspectiva longitudinal ("filme")	Dimensão revelada pelo tempo
Renda	Valor observado no momento da solicitação	Evolução, estabilidade, tendência, volatilidade e interrupções ao longo do tempo	Continuidade e estabilidade da geração de renda
Situação profissional	Ocupação atual	Histórico de vínculos, permanência nas atividades, transições e períodos de interrupção	Estabilidade ocupacional, empregabilidade e capacidade de recuperação
Relacionamento bancário	Banco, agência e conta informados na proposta	Tempo e continuidade do relacionamento, permanência em segmentos e intensidade de utilização	Maturidade e estabilidade do relacionamento financeiro
Trabalhador da *Gig Economy*	Estimativa pontual da renda	Evolução dos recebíveis, regularidade da atividade, sazonalidade, inatividade e recuperação	Continuidade da geração de renda, estabilidade operacional e resiliência

Fonte: Elaborada pelos autores.

A diferença entre as duas perspectivas torna-se mais concreta quando aplicada a informações aparentemente simples. Considere um candidato que informe, para recebimento dos recursos de um eventual empréstimo, uma conta mantida em agência ou segmento de atendimento diferenciado de determinada instituição financeira. Em uma abordagem transversal, banco, agência e conta constituem essencialmente informações cadastrais referentes ao momento da solicitação. Isoladamente, pouco revelam sobre a natureza do relacionamento financeiro existente.

Se houver informações históricas capazes de mostrar que essa conta é mantida continuamente há vários anos e que o cliente permaneceu durante esse período em determinado segmento de relacionamento, entretanto, torna-se possível construir atributos distintos daqueles contidos no cadastro original. Duração do vínculo, continuidade e permanência passam a representar aspectos da maturidade e estabilidade daquele relacionamento. Em certa medida, a engenharia de variáveis recupera, sob uma forma analítica, a lógica intuitiva do antigo "cliente desde" na folha de cheque: uma informação aparentemente simples que condensava anos de relacionamento bancário. O exemplo ilustra como o valor informacional pode não estar apenas no dado "conta bancária", mas na relação desse dado com o tempo.

O mesmo raciocínio assume especial relevância em situações nas quais a fotografia convencional representa de maneira limitada a realidade econômica do indivíduo. Considere um profissional da *Gig Economy* cuja renda decorra de entregas ou outros serviços realizados por meio de plataformas digitais. Uma estimativa pontual de renda mensal oferece uma representação particularmente frágil de uma atividade cuja remuneração pode variar entre semanas e meses. O valor observado em determinado período não permite distinguir, por exemplo, uma oscilação eventual de uma mudança persistente na atividade.

Quando se dispõe do histórico de recebíveis provenientes da plataforma, a engenharia de variáveis pode explorar sua dimensão temporal para construir atributos relacionados à regularidade da atividade, tendência dos recebimentos, sazonalidade, frequência e duração de períodos de inatividade e comportamento posterior às oscilações. O histórico deixa, assim, de constituir apenas uma sequência de valores e passa a permitir a representação de aspectos relacionados à continuidade da geração de renda, estabilidade operacional e capacidade de recuperação. Não se trata de converter recebíveis passados em garantia de pagamento futuro, mas de extrair da trajetória informações que uma estimativa pontual de renda não contém.

Os dois exemplos ilustram uma mesma propriedade da abordagem longitudinal. A informação adicional não decorre necessariamente da aquisição de uma nova base, mas pode emergir da transformação das relações existentes entre observações distribuídas no tempo. Isso diferencia a perspectiva longitudinal das sucessivas expansões do campo de observação discutidas na seção 2.2. Enquanto novas fontes permitem enxergar aspectos anteriormente indisponíveis, a abordagem longitudinal procura enxergar de outra maneira aquilo que já foi observado em diferentes momentos.

Essa distinção também permite situar a perspectiva longitudinal em relação à evolução dos algoritmos discutida na seção 2.3. Não se propõe substituir técnicas estatísticas, computacionais ou de aprendizado de máquina, mas ampliar a representação oferecida a essas técnicas. Os atributos longitudinais podem constituir variáveis de entrada para diferentes famílias de modelos; sua contribuição potencial encontra-se no conteúdo informacional que acrescentam à representação do indivíduo, e não na adoção de determinado algoritmo.

Sob essa ótica, a passagem da fotografia para o filme modifica a própria pergunta dirigida aos dados. Uma avaliação predominantemente transversal procura responder "como o indivíduo se apresenta hoje?". A perspectiva longitudinal acrescenta perguntas como "como chegou a essa condição?" e "o que sua trajetória permite conhecer que não estava visível na fotografia?". Essa mudança é particularmente relevante para a avaliação do potencial de crédito porque uma decisão tomada no presente procura inferir a capacidade de cumprimento de obrigações que se estenderão pelo futuro.

A perspectiva longitudinal pode, portanto, ser compreendida como uma forma adicional de enfrentar a insuficiência informacional discutida neste referencial. Historicamente, a avaliação do potencial de crédito evoluiu tornando observáveis novas dimensões da realidade e desenvolvendo métodos progressivamente mais capazes de processá-las. A oportunidade examinada neste relato acrescenta outra possibilidade: explorar o tempo como dimensão da engenharia de variáveis para transformar registros históricos em representações da trajetória econômico-financeira que não estavam explicitamente presentes nos dados originais. O tempo, nessa perspectiva, não apenas ordena as observações: contribui para tornar informacionalmente observável aquilo que somente existe na relação entre elas.

3 METODOLOGIA

Este relato técnico caracteriza-se como uma pesquisa aplicada, de natureza qualitativa e exploratória, estruturada segundo o protocolo para elaboração de relatos de produção técnica proposto por Biancolino, Kniess, Maccari e Rabechini Jr. (2012), e considerando também as diretrizes apresentadas por Martens, Pedron e Oliveira (2021) para artigos tecnológicos e relatos técnicos. Essas abordagens orientam a sistematização de problemas, oportunidades e intervenções originados em contextos profissionais, articulando experiência prática, fundamentação teórica e evidências empíricas.

O objeto do relato teve origem em experiência profissional diretamente vivenciada no desenvolvimento e aplicação de soluções analíticas voltadas à avaliação do potencial de crédito. A sistematização dessa experiência, entretanto, não se baseou exclusivamente na prática profissional. Para investigá-la e confrontá-la com outras perspectivas, o percurso metodológico articulou três fontes de evidência: a experiência aplicada em contexto organizacional, a literatura especializada e entrevistas com profissionais do mercado brasileiro.

A experiência profissional constituiu a primeira fonte de evidência e correspondeu à participação direta no desenvolvimento e aplicação de modelos de avaliação, motores de decisão e engenharia de variáveis em processos reais de avaliação do potencial de crédito. Essa experiência forneceu o contexto prático a partir do qual a oportunidade analisada no relato foi identificada e posteriormente sistematizada. Sua utilização como fonte de evidência decorre,

portanto, da participação profissional direta no fenômeno relatado, e não de observação externa ou exclusivamente documental.

A segunda fonte compreendeu informações secundárias provenientes da literatura acadêmica nacional e internacional e de documentos institucionais. A revisão contemplou trabalhos relacionados à assimetria informacional, avaliação de risco de crédito, evolução das fontes de informação, técnicas estatísticas e computacionais, aprendizado de máquina, engenharia de variáveis e utilização de dados alternativos. Essa etapa teve como finalidade proporcionar o enquadramento teórico da experiência profissional e estabelecer referências para sua posterior interpretação e confronto com as demais evidências.

A terceira fonte correspondeu a informações primárias obtidas por meio de entrevistas semiestruturadas com especialistas atuantes no mercado brasileiro de crédito, ciência de dados, prevenção a fraudes, cobrança, motores de decisão e desenvolvimento de modelos analíticos. Os participantes foram selecionados intencionalmente, priorizando profissionais com experiência em diferentes segmentos do ecossistema de crédito, de modo a reunir perspectivas complementares sobre os problemas e oportunidades investigados.

As entrevistas seguiram roteiro elaborado especificamente para o estudo, permitindo simultaneamente a comparação entre os depoimentos e a exploração de temas emergentes durante as conversas. Foram abordadas as limitações percebidas nos modelos e fontes de informação atualmente empregados, o papel dos dados alternativos, as perspectivas futuras para a avaliação do potencial de crédito e aspectos da realidade econômico-financeira considerados insuficientemente representados pelos mecanismos convencionais. A utilização de entrevistas semiestruturadas mostrou-se adequada por possibilitar aprofundamento analítico sem comprometer a comparabilidade entre os relatos (Kvale, 2011; Minayo, 2012).

As informações obtidas nas entrevistas foram submetidas à análise de conteúdo, buscando identificar padrões recorrentes, convergências, divergências e temas emergentes nos depoimentos (Bardin, 2016). Posteriormente, os resultados foram confrontados com as evidências provenientes da literatura e com a experiência profissional que originou o relato. Esse procedimento permitiu estabelecer uma triangulação entre prática aplicada, conhecimento teórico e percepção de profissionais de mercado, contribuindo para a robustez da interpretação qualitativa dos resultados (Flick, 2018).

A análise realizada possui delimitações importantes. O relato não compreendeu uma avaliação quantitativa retrospectiva destinada a isolar o efeito incremental das variáveis utilizadas na experiência profissional, tampouco realizou validação empírica de novas variáveis propostas a partir de sua sistematização. As entrevistas, por sua vez, tiveram natureza qualitativa e utilizaram seleção intencional dos participantes, não tendo como finalidade produzir inferências estatisticamente generalizáveis para o mercado brasileiro de crédito. Essas delimitações circunscrevem os resultados à sistematização da experiência aplicada e ao seu confronto qualitativo com evidências teóricas e profissionais.

Dessa forma, o percurso metodológico adotado pode ser sintetizado em três movimentos complementares: sistematização de uma experiência profissional aplicada, enquadramento e confronto com a literatura especializada e investigação qualitativa por meio de entrevistas com profissionais de mercado. A análise integrada dessas fontes constituiu a base para a apresentação e discussão dos resultados na seção seguinte.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

A experiência profissional que originou este relato ocorreu no desenvolvimento e aplicação de soluções analíticas voltadas à avaliação do potencial de crédito. Uma aplicação concreta teve início em 2014, no contexto da Geru, *fintech* brasileira criada para atuação no mercado de empréstimos pessoais *on-line*. Em registro sobre a trajetória da empresa, Cimatti

(2020) identifica a Geru como a primeira e, à época, maior fintech de empréstimos *on-line* do Brasil, tendo superado R$ 1 bilhão em empréstimos concedidos em seus primeiros cinco anos de operação. Cimatti (2020) destaca ainda o investimento em tecnologia e ciência de dados como elemento relevante de sua proposta de valor.

A natureza integralmente digital da operação tornava a avaliação remota, escalável e orientada por dados um componente central do modelo de negócio. Nesse ambiente, a experiência relatada envolveu participação direta no desenvolvimento e aplicação de modelos de avaliação e motores de decisão, bem como na concepção e engenharia de variáveis derivadas de diferentes fontes de informação. Entre os mecanismos empregados esteve a construção de atributos a partir de observações históricas de determinados fenômenos, incorporados aos processos de avaliação utilizados pela *fintech*.

Essa aplicação constitui a primeira evidência relevante deste relato: a exploração longitudinal de informações históricas mostrou-se operacionalmente aplicável em um processo real de avaliação do potencial de crédito. A experiência, entretanto, não permite, isoladamente, estabelecer a contribuição incremental específica desses atributos para o desempenho dos modelos, uma vez que sua utilização ocorreu integrada a outras variáveis e mecanismos de decisão e não foi objeto, neste estudo, de análise retrospectiva destinada a isolar seus efeitos.

Para verificar se o problema identificado na experiência profissional encontrava correspondência em outros contextos, foram analisadas as entrevistas realizadas com especialistas do mercado brasileiro. Os depoimentos apresentaram elevada convergência quanto à permanência de lacunas relevantes na representação da realidade econômico-financeira dos consumidores, apesar da ampliação das fontes de informação e da evolução das técnicas analíticas. A Tabela 4 sintetiza os principais temas identificados.

Tabela 4 – Síntese dos principais achados obtidos nas entrevistas com especialistas

Tema analisado	Principais evidências observadas	Convergência
Limitações dos dados tradicionais	Fontes maduras e amplamente utilizadas, porém insuficientes para representar integralmente a realidade econômica dos indivíduos	Alta
Continuidade e estabilidade de renda	A estimação da renda atual evoluiu, mas sua continuidade e estabilidade futura permanecem uma lacuna relevante	Muito alta
Consumidores com baixa densidade informacional	Trabalhadores informais, autônomos e indivíduos com pouco histórico financeiro permanecem entre os mais difíceis de avaliar	Muito alta
Dados alternativos	Considerados promissores como mecanismo complementar de ampliação informacional	Muito alta
Engenharia de variáveis	O valor não reside apenas na fonte bruta, mas na capacidade de transformar dados em atributos aderentes ao problema	Alta
Necessidade de validação empírica	Eventuais ganhos precisam ser comprovados mediante testes quantitativos	Muito alta
Evolução da análise de risco	A ampliação da suficiência informacional foi percebida como relevante para avanços futuros, para além da adoção de algoritmos mais sofisticados	Muito alta

Fonte: Dados da pesquisa

Esse achado sustenta, sem demonstrar quantitativamente, a pertinência da perspectiva longitudinal apresentada no referencial teórico. A preocupação manifestada pelos especialistas não se concentrou apenas na obtenção de uma estimativa mais precisa da condição atual, mas também na dificuldade de representar características que se manifestam ao longo do tempo. Nesse sentido, as entrevistas oferecem evidência de que o problema que motivou a experiência profissional não se restringe ao contexto organizacional em que ela se originou.

A dificuldade mostrou-se especialmente relevante para trabalhadores autônomos, profissionais da *Gig Economy*, microempreendedores e consumidores com baixa densidade informacional, apontados entre os segmentos mais difíceis de avaliar pelos mecanismos convencionais. Para esses públicos, a ausência ou escassez de informações tradicionalmente utilizadas não implica necessariamente ausência de capacidade econômica, mas pode limitar sua adequada representação nos processos de avaliação. Esse resultado converge com estudos que destacam o potencial de informações alternativas para complementar a avaliação de indivíduos caracterizados por informalidade econômica ou histórico financeiro limitado (Jagtiani & Lemieux, 2019; Hlongwane et al., 2024).

Outro resultado recorrente nas entrevistas refere-se à engenharia de variáveis. Os especialistas indicaram que a ampliação da quantidade de dados disponíveis não produz automaticamente melhores decisões: seleção, interpretação e transformação das informações permanecem relevantes para a construção de atributos aderentes ao fenômeno avaliado. Essa percepção é consistente com Hand e Henley (1997), Lessmann et al. (2015) e Louzada et al. (2016), que, sob diferentes perspectivas, demonstram que o desempenho dos modelos não depende exclusivamente da sofisticação dos algoritmos.

Os depoimentos também qualificam a oportunidade examinada ao evidenciarem dificuldades relacionadas à densidade informacional, operacionalização de novas fontes, qualidade dos dados e seleção das informações efetivamente relevantes entre volumes crescentes de registros. Os resultados sugerem, portanto, que maior disponibilidade de dados não deve ser interpretada automaticamente como maior suficiência informacional. Essa constatação converge com o problema de representação discutido no referencial e ajuda a delimitar o espaço no qual abordagens complementares de engenharia de variáveis podem ser investigadas.

A triangulação das três fontes de evidência permite estabelecer, assim, diferentes níveis de sustentação para a oportunidade analisada. A experiência profissional demonstra sua aplicabilidade operacional; as entrevistas indicam que as lacunas informacionais que a motivaram são reconhecidas por profissionais atuantes em diferentes contextos; e a literatura fornece respaldo à importância da qualidade das variáveis, das fontes e da representação dos fenômenos na avaliação de crédito. Em conjunto, essas evidências sustentam a pertinência de investigar a perspectiva longitudinal como mecanismo complementar de representação da realidade econômico-financeira.

Essa convergência, contudo, deve ser interpretada dentro dos limites do estudo. A experiência relatada não foi submetida a uma análise retrospectiva destinada a isolar quantitativamente a contribuição incremental dos atributos longitudinais utilizados, e as entrevistas produziram evidências qualitativas, não generalizáveis estatisticamente. Os próprios especialistas ressaltaram a necessidade de validação quantitativa de eventuais ganhos associados a novas fontes ou atributos. Portanto, os resultados não permitem afirmar que variáveis longitudinais produzirão necessariamente maior capacidade discriminatória em qualquer contexto; indicam que a abordagem é profissionalmente aplicável, teoricamente consistente e aderente a lacunas reconhecidas pelos especialistas, justificando sua investigação empírica em aplicações específicas.

4.1 Contribuição e aprendizado profissional

O principal aprendizado decorrente da experiência e de sua sistematização consiste em reconhecer que a busca por maior suficiência informacional pode incluir, entre suas alternativas, a exploração mais estruturada dos históricos já disponíveis nas organizações. Para profissionais de crédito, risco e ciência de dados, isso acrescenta uma questão prática às decisões sobre

aquisição de novas fontes ou desenvolvimento de modelos: que informação relevante ainda pode ser extraída dos dados históricos disponíveis por meio da engenharia de variáveis?

Essa questão apresenta potencial de transferência para organizações que disponham de séries de observações sobre seus clientes, mas não pressupõe que todo histórico produzirá atributos úteis. A pertinência da abordagem dependerá da natureza do fenômeno analisado, da qualidade e extensão das observações, da consistência econômica das variáveis construídas e, principalmente, da demonstração empírica de sua contribuição incremental para a decisão.

A aplicação pode ser particularmente promissora em contextos nos quais a condição econômico-financeira é pouco representada por informações pontuais, como ocorre com determinados trabalhadores autônomos, profissionais da *Gig Economy* e outros públicos com renda irregular ou baixa densidade de histórico financeiro convencional. Nesses casos, históricos de atividade ou recebimentos constituem matéria-prima para investigar empiricamente atributos complementares aos mecanismos tradicionais de avaliação.

O aprendizado profissional socializado neste relato não consiste, portanto, em recomendar a substituição de fontes, modelos ou técnicas existentes por uma abordagem longitudinal, mas em incorporá-la ao conjunto de alternativas disponíveis para a engenharia de variáveis. Sua contribuição efetiva deverá ser estabelecida caso a caso, mediante formulação de atributos economicamente interpretáveis e posterior validação quantitativa.

5 CONSIDERAÇÕES FINAIS

Este relato técnico teve como objetivo sistematizar e analisar uma experiência profissional aplicada relacionada à utilização da dimensão temporal na engenharia de variáveis e reunir evidências teóricas e empíricas sobre sua pertinência para ampliar a suficiência informacional nos processos de avaliação do potencial de crédito. A articulação entre experiência profissional, literatura especializada e entrevistas com profissionais de mercado permitiu atingir esse objetivo, ao evidenciar a aplicabilidade da abordagem em contexto organizacional e sua aderência a lacunas informacionais reconhecidas tanto pela literatura quanto pelos especialistas consultados.

A principal contribuição do relato consiste em sistematizar uma perspectiva complementar para o aprimoramento da avaliação do potencial de crédito: antes de depender exclusivamente da incorporação de novas fontes de dados ou de algoritmos progressivamente mais sofisticados, organizações podem investigar que informação adicional pode ser construída a partir dos históricos de que já dispõem. A perspectiva longitudinal amplia, assim, o espaço de atuação da engenharia de variáveis ao incorporar a dimensão temporal à representação da realidade econômico-financeira dos indivíduos.

Essa contribuição não implica substituir fontes tradicionais, dados alternativos, técnicas estatísticas, métodos computacionais ou modelos de aprendizado de máquina. Ao contrário, a abordagem proposta é complementar a esses recursos e pode fornecer atributos adicionais para diferentes famílias de modelos. Sua relevância potencial reside particularmente em situações nas quais informações pontuais representam de maneira incompleta aspectos economicamente relevantes do candidato, cabendo à validação empírica estabelecer a contribuição efetiva desses atributos em cada aplicação.

Para organizações que enfrentem problemas semelhantes, o relato sugere um caminho prático: antes de buscar exclusivamente novas bases, examinar sistematicamente os históricos já disponíveis e avaliar se eles contêm relações temporalmente estruturadas que possam ser convertidas em atributos economicamente interpretáveis. Essa análise pode ser particularmente pertinente em contextos de renda irregular, informalidade ou baixa densidade de informações convencionais. Os atributos desenvolvidos devem, entretanto, ser submetidos aos mesmos

critérios de qualidade, estabilidade e validação estatística e computacional aplicáveis às demais variáveis utilizadas nos processos de decisão.

O estudo apresenta limitações que delimitam o alcance de suas conclusões. Embora a experiência relatada demonstre a viabilidade de construção e utilização operacional de atributos longitudinais, não foi realizada análise quantitativa retrospectiva destinada a isolar sua contribuição incremental para o desempenho dos modelos empregados. As entrevistas, por sua vez, tiveram natureza qualitativa e seleção intencional dos participantes, devendo seus resultados ser interpretados como evidências de convergência profissional, e não como resultados estatisticamente generalizáveis para o mercado brasileiro de crédito. Consequentemente, o relato não permite afirmar que atributos longitudinais produzirão necessariamente ganhos de discriminação ou desempenho em qualquer contexto.

Essas limitações apontam diretamente para a agenda futura. O próximo passo consiste no desenvolvimento e na validação quantitativa de famílias de variáveis longitudinais em diferentes contextos de avaliação do potencial de crédito, comparando modelos com e sem esses atributos e mensurando sua contribuição incremental, estabilidade temporal e capacidade discriminatória. Particular interesse reside em aplicações envolvendo informações históricas de baixo custo operacional e públicos cuja realidade econômica seja insuficientemente representada pelas fontes convencionais, como trabalhadores autônomos e profissionais da *Gig Economy*.

Novos estudos e relatos técnicos poderão ainda investigar a transferibilidade da abordagem para diferentes contextos organizacionais e fontes de informação. Dessa forma, a experiência profissional sistematizada neste relato pode constituir um ponto de partida para a construção progressiva de evidências sobre quando, em quais contextos e em que magnitude a incorporação da dimensão temporal à engenharia de variáveis efetivamente melhora a avaliação do potencial de crédito.

REFERÊNCIAS

Banco Central do Brasil. (2024b). *Cadastro Positivo.* https://www.bcb.gov.br/meubc/faqs/s/cadastro-positivo.

Banco Central do Brasil. (2024c). *Sistema de Informações de Créditos (SCR).* https://www.bcb.gov.br/estabilidadefinanceira/scr.

Banco Central do Brasil. (2024d). *Relatório de estabilidade financeira.* https://www.bcb.gov.br/content/publicacoes/ref/202410/RELESTAB202410-refPub.pdf.

Banco Central do Brasil. (2024e). *O que é o Pix?* https://www.bcb.gov.br/estabilidadefinanceira/pix.

Bardin, L. (2016). *Análise de conteúdo.* Edições 70.

Biancolino, C. A., Kniess, C. T., Maccari, E. A., & Rabechini Jr, R. (2012). Protocolo para elaboração de relatos de produção técnica. *Revista de Gestão e Projetos, 3(2), 294–307. https://doi.org/10.5585/gep.v3i2.121*.

Cimatti, S. (2020). *Prosperidade, felicidade, fintechs e a pandemia.* Scortecci.

Febraban. (2024). *Pesquisa Febraban de tecnologia bancária 2024 (Vol. 1).* Deloitte & Federação Brasileira de Bancos.

Flick, U. (2018). *An introduction to qualitative research* (6th ed.). Sage.

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). BigTech and the changing structure of financial intermediation. *Economic Policy, 34(100), 761–799*.

Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 160(3), 523–541. http://www.jstor.org/stable/2983268*.

Hlongwane, R., Ramaboa, K. K., & Mongwe, W. (2024). Enhancing credit scoring accuracy with a comprehensive evaluation of alternative data. *PLoS ONE, 19(5), e0303566. https://doi.org/10.1371/journal.pone.0303566*.

Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of big data. *Yale Journal of Law and Technology, 18(1), 148–216.*

Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Financial Management, 48(4), 1009–1029*.

Kvale, S. (2011). Doing interviews. *SAGE Publications, https://doi.org/10.4135/9781849208963*.

Lawrence, D. B. (1987). *Risco & recompensa: O negócio de crédito ao consumidor, uma doutrina instrumental.* (D. McKey, Trad.; C. H. M. de Alvarenga, Rev. téc. e introd.). Bandeirante.

Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247(1), 124–136*.

Louzada, F., Ara, A., & Fernandes, G. B. (2016). Classification methods applied to credit scoring: Systematic review and overall comparison. *Surveys in Operations Research and Management Science, 21(2), 117–134. https://doi.org/10.1016/j.sorms.2016.10.001*.

Martens, C. D., Pedron, C. D., & Oliveira, J. C. (2021). Diretrizes para elaboração de artigos tecnológicos, artigos aplicados ou relatos técnicos de produção com ênfase profissional. *Revista Inovação, Projetos E Tecnologias, 9(2), 143–147. https://doi.org/10.5585/iptec.v9i2.21117*.

Minayo, M. C. (2012). Análise qualitativa: teoria, passos e fidedignidade. *Ciência & Saúde Coletiva, 17(3) 621–626*.

Open Finance Brasil. (2022). *Conheça o Open Finance.* https://openfinancebrasil.org.br/conheca-o-open-finance.

Souto, G. A. (2019). Cadastro Positivo: A solução para o combate à assimetria informacional no setor bancário brasileiro? *Revista da Procuradoria-Geral do Banco Central, 13(1), 75–88. https://doi.org/10.58766/rpgbcb.v13i1.1016*.

Sprague, J. R. (1946). *O romance do crédito.* (M. Rangel, Trad.) Irmãos Di Giorgio & Cia.

Stiglitz, J. E., & Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American Economic Review, 71(3), 393–410. http://www.jstor.org/stable/1802787*.

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2017). *Credit scoring and its applications.* SIAM Society for Industrial and Applied Mathematics.