Revista Ibero-Americana de Humanidades, Ciências e Educação — REASE

OPEN ACCESS

doi.org/10.51891/rease.v12i11.33737

LEAD SCORING COM REGRESSÃO LOGÍSTICA APLICADO À PROSPECÇÃO B2B: UM MODELO DE PRIORIZAÇÃO DE LEADS EM OPERAÇÃO COMERCIAL REAL

LEAD SCORING WITH LOGISTIC REGRESSION APPLIED TO B2B PROSPECTING: A LEAD PRIORITIZATION MODEL IN A REAL COMMERCIAL OPERATION

LEAD SCORING CON REGRESIÓN LOGÍSTICA APLICADO A LA PROSPECCIÓN B2B: UN MODELO DE PRIORIZACIÓN DE LEADS EN UNA OPERACIÓN COMERCIAL REAL

José Lucas Marques Silva[1]
Ítalo César Carvalho Lima[2]

RESUMO: Em operações comerciais de pequeno porte que prospectam outras organizações, a ordem de abordagem dos leads costuma depender do julgamento do operador. Este estudo teve como objetivo propor e validar um modelo de lead scoring baseado em regressão logística em uma operação real de prospecção B2B, comparando-o com a priorização praticada. Adotou-se delineamento misto sequencial exploratório, em estudo de caso único: um memorando do pesquisador, registrado antes do ajuste do modelo, mapeou os critérios de priorização, e uma base anonimizada de 89 leads, com 15 avanços no funil, foi usada para estimar o modelo, validado por correção de otimismo com bootstrap, e compará-lo com a ordem histórica de abordagem segundo regra de decisão fixada previamente. Dos oito critérios identificados, apenas o segmento de atuação estava registrado. O segmento considerado de melhor encaixe teve a menor taxa de avanço (7,1% contra 21,3%), sem significância estatística. O score elevou em 6,3 pontos percentuais a taxa de avanço nos primeiros 20% da fila, mas o intervalo de confiança incluiu o zero, e a hipótese foi classificada como inconclusiva. Conclui-se que a priorização preditiva depende do registro sistemático dos atributos que o operador já observa.

Palavras-chave: Lead scoring. Regressão logística. Prospecção B2B.

ABSTRACT: In small commercial operations that prospect other organizations, the order in which leads are approached usually depends on the operator's judgment. This study aimed to propose and validate a lead scoring model based on logistic regression in a real B2B prospecting operation, comparing it with the prioritization in use. A sequential exploratory mixed-methods design was adopted in a single case study: a researcher memo, recorded before the model was fitted, mapped the prioritization criteria, and an anonymized base of 89 leads, with 15 funnel advances, was used to estimate the model, validated by bootstrap optimism correction, and to compare it with the historical approach order under a decision rule set in advance. Of the eight criteria identified, only the business segment was recorded. The segment considered the best fit had the lowest advance rate (7.1% against 21.3%), without statistical significance. The score raised the advance rate in the first 20% of the queue by 6.3 percentage points, but the confidence interval included zero, and the hypothesis was classified as inconclusive. Predictive prioritization depends on systematically recording the attributes the operator already observes.

Keywords: Lead scoring. Logistic regression. B2B prospecting.

[1] Graduando em Engenharia de Software. Instituto de Ensino Superior iCEV.
[2] MBA em Marketing, Branding e Growth pela Pontifícia Universidade Católica do Rio Grande do Sul (PUC-RS). Orientador. Instituto de Ensino Superior iCEV.

RESUMEN: En pequeñas operaciones comerciales que prospectan otras organizaciones, el orden de abordaje de los leads suele depender del juicio del operador. Este estudio tuvo como objetivo proponer y validar un modelo de lead scoring basado en regresión logística en una operación real de prospección B2B, comparándolo con la priorización practicada. Se adoptó un diseño mixto secuencial exploratorio, en un estudio de caso único: un memorando del investigador, registrado antes del ajuste del modelo, mapeó los criterios de priorización, y una base anonimizada de 89 leads, con 15 avances en el embudo, se utilizó para estimar el modelo, validado por corrección de optimismo con bootstrap, y compararlo con el orden histórico de abordaje según una regla fijada previamente. De los ocho criterios identificados, solo el segmento de actuación estaba registrado. El segmento considerado de mejor encaje tuvo la menor tasa de avance (7,1% frente a 21,3%), sin significancia estadística. El score elevó en 6,3 puntos porcentuales la tasa de avance en el primer 20% de la fila, pero el intervalo de confianza incluyó el cero, y la hipótesis se clasificó como inconclusa. Se concluye que la priorización predictiva depende del registro sistemático de los atributos que el operador ya observa.

Palabras clave: Lead scoring. Regresión logística. Prospección B2B.

1 INTRODUÇÃO

A digitalização dos processos comerciais ampliou a capacidade de identificar potenciais clientes em mercados *business-to-business* (B2B). Ferramentas de prospecção, bases públicas de dados cadastrais e canais digitais de contato permitem que mesmo uma operação comercial de pequeno porte mapeie centenas de organizações aderentes ao seu perfil de cliente. O gargalo, contudo, deslocou-se: deixou de estar na identificação de leads e passou a estar na capacidade de atendê-los. Cada contato exige pesquisa prévia, abordagem personalizada e acompanhamento ao longo de uma cadência, atividades cujo custo em tempo cresce aproximadamente de forma linear com o número de leads trabalhados. Em operações que não dispõem de equipe dedicada de pré-vendas, esse tempo é o recurso mais escasso do processo comercial.

Diante dessa restrição, a ordem em que os leads são abordados deixa de ser um detalhe operacional e passa a condicionar o resultado comercial do período. Abordar primeiro os leads com maior probabilidade de avanço significa obter respostas e reuniões mais cedo, com o mesmo esforço empregado. A prática corrente em operações de pequeno porte, entretanto, costuma apoiar-se no julgamento do próprio operador, que classifica os leads em níveis de prioridade segundo critérios como segmento e porte, por vezes com apoio de ferramentas de triagem. Tal priorização é simples de executar, porém não é confrontada com aquilo que o histórico de interações já revelou sobre quais características antecedem o avanço do lead no funil.

A literatura de *lead scoring* sistematiza essa decisão. Wu M, et al. (2024), em revisão sistemática de 44 publicações, distinguem os modelos tradicionais, baseados em julgamento humano, dos modelos preditivos, construídos a partir de dados históricos, e concluem que os segundos apresentam desempenho superior em vendas, ainda que a um custo de implementação

mais elevado. A aplicação de técnicas estatísticas à conversão de leads B2B não é recente (MONAT JP, 2011; D'HAEN J e VAN DEN POEL D, 2013), e González-Flores L, et al. (2025) construíram um modelo de *lead scoring* em empresa de *software* B2B a partir da comparação entre quinze algoritmos de classificação.

No Brasil, essa agenda permanece pouco explorada em termos empíricos. Coda RC e Castro GHC (2019), ao mapearem a produção científica brasileira em marketing B2B entre 2008 e 2018, identificaram presença reduzida de temas ligados à análise de dados, e Barros PM e Souza CA (2025) revisaram a contribuição da inteligência artificial nos relacionamentos B2B sem construir nem validar modelo a partir dos dados de uma operação. A produção nacional localizada inclui classificadores de leads construídos com dados reais de empresas (OLIVEIRA AVM e SILVA PHL, 2022; REIS VG, 2025), mas avalia seu desempenho sem confrontar a qualificação produzida pelo modelo com o critério que a operação já empregava. É essa comparação, com regras de decisão definidas antes da análise, que este estudo realiza.

A relevância dessa lacuna extrapola o interesse acadêmico. Segundo o Sebrae (2026), o Brasil conta com 24 milhões de micro e pequenas empresas ativas, que representam 95% do total de empresas do país e respondem por 26,5% do Produto Interno Bruto. Em sua maioria, trata-se de organizações que conduzem a prospecção comercial sem equipe especializada e sem instrumentos formais de apoio à decisão. Para esse perfil, um modelo de priorização construído sobre dados da própria operação e interpretável por quem o utiliza representa ganho de produtividade obtido sem ampliação de estrutura.

Diante do exposto, formula-se o seguinte problema de pesquisa: como um modelo de *lead scoring* baseado em regressão logística pode apoiar a priorização de leads em um processo de prospecção B2B, aumentando a taxa de avanço no funil comercial em relação à priorização atualmente praticada? O objetivo geral consiste em propor e validar esse modelo a partir dos dados históricos de uma operação comercial real. Os objetivos específicos são: (i) mapear os critérios utilizados, implícita ou explicitamente, para selecionar e priorizar leads; (ii) estruturar o histórico de abordagem e de avanço no funil; (iii) desenvolver um modelo que ordene os leads por probabilidade de avanço; e (iv) comparar o modelo com a priorização praticada, pela taxa de avanço e pelo esforço até o avanço, expresso em número de leads abordados até o primeiro avanço e até metade dos avanços.

A hipótese é a de que a priorização por *score* supera a priorização praticada, por ponderar os atributos dos leads com pesos estimados a partir das respostas que a própria operação obteve,

enquanto a priorização praticada se apoia em julgamento qualitativo, ainda que assistido por ferramenta, não confrontado com os desfechos observados. A regra que decide sobre a hipótese foi fixada antes da análise e admite três classificações, sustentada, inconclusiva ou não sustentada, conforme a posição do intervalo de confiança da diferença entre as duas ordenações em relação a zero.

2 REFERENCIAL TEÓRICO

O processo comercial B2B é usualmente representado como um funil que se estreita a cada etapa, com perda de volume entre os estágios e consumo de tempo e esforço em cada um deles. A geração de leads não garante seu aproveitamento: Sabnis G, et al. (2013) constatam que cerca de 70% dos leads gerados pelo marketing não são trabalhados pelos vendedores, fenômeno que denominam buraco negro dos leads, e, em pesquisa com 461 vendedores de quatro empresas, identificam a qualidade da pré-qualificação entre os fatores que explicam o acompanhamento. Em operações de prospecção ativa (*outbound*), como a estudada, não há estágio de marketing que preceda o contato: o próprio operador identifica, seleciona e prioriza os leads segundo critérios como setor ou porte. É nesse momento, o da decisão sobre quais leads abordar e em que ordem, que o modelo proposto atua.

Lead scoring é a atribuição de uma pontuação a cada lead, destinada a expressar sua probabilidade de conversão e, por consequência, a ordenar o esforço comercial. Wu M, et al. (2024) organizam o campo em duas famílias. Os modelos tradicionais atribuem pontos a atributos do lead com base no julgamento de profissionais de marketing e vendas; oferecem simplicidade, mas dependem da experiência de quem define os pesos e não se ajustam automaticamente quando o padrão de conversão se altera. A priorização por níveis definidos pelo operador constitui uma forma de *lead scoring* tradicional, ainda que expressa em categorias. Os modelos preditivos estimam a pontuação a partir de dados históricos e apresentam desempenho superior em vendas, com destaque para os algoritmos de classificação e a regressão logística.

D'Haen J e Van den Poel D (2013) propuseram um *framework* de três fases para a prospecção B2B, no qual a predição se refina à medida que se acumula retorno sobre os leads não convertidos, com o emprego, na segunda fase, de regressão logística, árvores de decisão e redes neurais. O *framework* evidencia um ponto crítico para operações de pequeno porte: o modelo preditivo depende de histórico, e o histórico só existe depois que a operação já

prospectou. Essa é a perspectiva adotada neste estudo, cuja base provém de uma operação em funcionamento, com histórico ainda limitado. Em contribuição recente, González-Flores L, et al. (2025) compararam quinze algoritmos em uma empresa de *software* B2B; o *Gradient Boosting* obteve o melhor desempenho, e o modelo superou o *scoring* manual até então praticado pela empresa.

A regressão logística modela a probabilidade de ocorrência de um evento binário, como o avanço ou não de um lead no funil, por meio do logaritmo da razão de chances, expresso como combinação linear das variáveis explicativas (HOSMER DW, et al., 2013), conforme a Equação 1:

$$\ln[p / (1 - p)] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \quad (1)$$

em que p é a probabilidade de avanço, x_1 a x_k são as características do lead e β_0 a β_k, os coeficientes estimados. A saída é uma probabilidade entre 0 e 1, utilizável diretamente como *score*, e o termo $\exp(\beta_i)$ corresponde à razão de chances associada à variável x_i, o que permite a quem utiliza o modelo identificar quanto cada característica pesa na recomendação (HOSMER DW, et al., 2013). Essa legibilidade sustenta a posição de Rudin C (2019), para quem decisões que exigem confiança devem empregar modelos interpretáveis desde a origem, e não modelos opacos acompanhados de explicações posteriores. Este estudo adota essa posição e assume a possível perda de acurácia como limitação declarada.

A principal restrição da técnica em bases pequenas refere-se ao número de eventos por variável (EPV), razão entre o número de casos da categoria menos frequente do desfecho e o número de preditoras. Peduzzi P, et al. (1996) verificaram, em simulação, que valores inferiores a 10 produzem coeficientes enviesados e intervalos com cobertura inadequada, enquanto Vittinghoff E e McCulloch CE (2007) sustentam que a regra pode ser relaxada em circunstâncias específicas. Na prática, o número de variáveis do modelo fica limitado pela quantidade de leads que efetivamente avançaram, e não pelo volume de informação disponível sobre cada lead.

A produção nacional recente sobre qualificação de leads propõe ou testa métodos de classificação, como a lógica *fuzzy* aplicada aos leads de uma emissora de televisão (LIMA AFO, et al., 2024), a predição de links sobre a base de 5.046 leads de uma organização que vende cursos (BRUGALLI GP, et al., 2022) e classificadores de aprendizado de máquina treinados com dados reais: uma rede neural sobre os leads de uma empresa de projetos residenciais, com acurácia de 97,9% (RABELO AR, 2019), árvores de decisão sobre os leads de um funil B2B extraídos de um

CRM, com acurácia de 94,7% (OLIVEIRA AVM e SILVA PHL, 2022), e o modelo *random forest* sobre 72 conversas de uma concessionária de veículos, com acurácia média de 86% em validação cruzada (REIS VG, 2025). Em busca complementar, Costa OS e Gouveia LB (2021) aplicaram regressão logística aos leads de um polo de uma universidade do Sul do Brasil e relataram acurácia de 100%. Esses trabalhos, em geral, avaliam os modelos sobretudo pela acurácia, medida que, em bases com poucos convertidos, pode ser elevada mesmo quando o modelo erra a maior parte deles, ou deixam de confrontá-los com desfechos observados. Nenhum compara a ordenação produzida pelo modelo com a priorização que a operação já praticava, e a busca descrita nos métodos não localizou estudo nacional que realize essa comparação.

3 MÉTODOS

Trata-se de pesquisa aplicada, de abordagem mista em delineamento sequencial exploratório, no qual uma etapa qualitativa inicial informa a etapa quantitativa que a sucede (CRESWELL JW e CRESWELL JD, 2021). A etapa qualitativa é exploratória e a quantitativa, descritiva. O procedimento é o estudo de caso único, com pesquisa documental sobre registros produzidos na rotina da operação e pesquisa *ex post facto*, já que os dados se referem a fatos ocorridos sem manipulação de variáveis pelo pesquisador (GIL AC, 2022). Os dados são observacionais e a seleção dos leads abordados não foi aleatória, de modo que os coeficientes do modelo expressam associações úteis à predição, e não efeitos causais.

O primeiro autor integra a operação estudada e participa da execução da prospecção, na condição de praticante-observador, o que Brannick T e Coghlan D (2007) denominam pesquisa de *insider*. Essa condição assegura acesso integral aos registros, mas cria o risco de que o pesquisador avalie favoravelmente a própria proposta. Para mitigá-lo, adotaram-se três salvaguardas: o memorando com os critérios de priorização foi redigido antes do ajuste do modelo; as variáveis, as métricas e o critério de comparação foram fixados antes do contato com os resultados; e a linha de base corresponde ao registro histórico da operação, e não a uma reconstrução posterior.

O estudo foi realizado na operação comercial de uma empresa de pequeno porte que presta serviços a outras organizações (B2B), que conduz a prospecção de forma ativa e não dispõe de equipe dedicada de pré-vendas. Os leads são localizados em buscadores de mapas e em redes sociais e submetidos a uma triagem assistida por ferramenta de inteligência artificial

generativa, que, a partir do perfil público de cada organização, registra o segmento de atuação e atribui um nível de prioridade, revisado pelo operador. A abordagem segue esses níveis, dos mais altos para os mais baixos, e, dentro de cada nível, a ordem da triagem. O contato é feito exclusivamente por aplicativo de mensagens, com um primeiro contato e até dois contatos de acompanhamento, espaçados em três dias úteis. Em atenção ao sigilo comercial, a organização é caracterizada apenas em termos genéricos.

A unidade de análise é o lead, entendido como a organização prospectada e abordada. A amostra reúne os leads abordados entre 17 de agosto e 23 de outubro de 2026, com desfecho observado até a extração dos dados, em 16 de novembro de 2026, e totaliza 89 leads, dos quais 15 avançaram no funil. Adotou-se uma janela de observação, e não a conclusão formal da cadência, porque a prospecção foi interrompida antes que a maior parte dos leads recebesse todos os contatos previstos; a janela assegura a cada lead pelo menos três semanas entre o primeiro contato e a extração. Foram excluídos os leads não abordados, os de número de contato inválido e os abordados por outro integrante da operação.

A etapa qualitativa mapeou os critérios de seleção e priorização de leads a partir de três fontes: os registros de triagem e de contato, os modelos de mensagem da cadência e um memorando analítico registrado pelo primeiro autor, em resposta a um roteiro de perguntas abertas, antes do ajuste do modelo. Cada critério foi classificado segundo três atributos: se é explícito ou implícito; se é conhecido antes ou somente depois do primeiro contato; e se está registrado de forma estruturada. O quadro resultante forneceu as variáveis candidatas ao modelo e explicitou a regra da ordenação de referência.

A base foi extraída dos registros operacionais da prospecção, estruturada com um registro por lead e anonimizada antes de qualquer análise. A variável-alvo, o avanço no funil, assume valor 1 quando o lead respondeu com abertura para conversa ou teve reunião agendada, e o quando não houve resposta até a extração ou houve recusa. Respostas automáticas foram tratadas como ausência de resposta. Considerou-se abertura a resposta que dá continuidade à conversa sem declarar ausência de necessidade ou de interesse, e recusa a que declara, direta ou indiretamente, não precisar da solução ou não ter interesse. Como duas respostas tiveram classificação ambígua, os resultados foram recalculados com a classificação alternativa, como análise de sensibilidade.

As preditoras foram restritas a informações conhecidas antes do primeiro contato, porque o uso pretendido do modelo é ordenar leads ainda não abordados; variáveis registradas

durante a cadência, como o número de tentativas, configurariam vazamento, isto é, informação sobre a variável-alvo que não estaria legitimamente disponível no momento da decisão (KAUFMAN S, et al., 2012). Com 15 eventos, o patamar de 10 eventos por variável (PEDUZZI P, et al., 1996) admite um único parâmetro. O nível de prioridade atribuído na triagem não integra as preditoras, porque orienta a ordenação de referência e porque a prioridade final, revisada pelo operador, não ficou registrada de forma estruturada; a ordenação de referência foi, por isso, reconstituída pela data efetiva do primeiro contato. O modelo foi especificado com uma única preditora, selecionada a priori: o segmento de atuação, registrado antes do contato e agrupado em duas categorias definidas no memorando, o segmento que o operador identifica como de maior volume de demanda e os demais. A distribuição dos desfechos por segmento foi visualizada durante a preparação da base; o agrupamento, contudo, reproduz o critério registrado no memorando, e não aquela distribuição.

O modelo foi estimado por regressão logística binária, por máxima verossimilhança, na forma da Equação 1 (HOSMER DW, et al., 2013), com os coeficientes reportados como razões de chances e intervalos de confiança de 95%. A significância global foi avaliada pelo teste da razão de verossimilhança, o ajuste pelo teste de Hosmer-Lemeshow, e verificou-se a ocorrência de separação completa ou quase completa. A escolha da regressão logística, e não de algoritmos de maior capacidade preditiva, apoia-se na interpretabilidade, na adequação ao tamanho da amostra e na aplicabilidade para um operador sem formação em ciência de dados. A análise foi implementada em Python, com as bibliotecas statsmodels (SEABOLD S e PERKTOLD J, 2010) e scikit-learn (PEDREGOSA F, et al., 2011); o código, desenvolvido com apoio de ferramenta de inteligência artificial generativa sob especificação e conferência do primeiro autor, está disponível em repositório público (https://github.com/Dante-teslq/TCC-Lead-Scoring-V1.0).

A validação interna foi feita por correção de otimismo com *bootstrap*, procedimento que Steyerberg EW, et al. (2001) identificaram como o mais eficiente para modelos de regressão logística, enquanto a separação da amostra em treino e teste se mostrou ineficiente. Mede-se o desempenho aparente, do modelo ajustado e avaliado na amostra completa; em cada uma de 2.000 reamostras com reposição, o modelo é reajustado e avaliado na reamostra e na amostra original, e a diferença entre as duas medidas é o otimismo daquela reamostra; o desempenho corrigido é o aparente menos a média dos otimismos. A discriminação foi medida pela área sob a curva ROC (AUC), em que 0,5 equivale a uma ordenação aleatória e valores entre 0,7 e 0,8 indicam discriminação aceitável (HOSMER DW, et al., 2013). Métricas que dependem de ponto

de corte, como a acurácia, não foram adotadas como principais, porque o uso pretendido é ordenar leads e porque, com classes desbalanceadas, a acurácia pode ser elevada mesmo para um classificador que atribui todos os casos à classe majoritária.

A ordenação de referência corresponde à ordem em que os leads foram efetivamente abordados, e a ordenação proposta dispõe os mesmos leads em ordem decrescente de *score*. As duas foram confrontadas pela taxa de avanço acumulada nos primeiros 10%, 20% e 30% da fila, com o recorte de 20% fixado antes da análise como métrica principal, e pelo esforço até o avanço, medido pelo número de leads abordados até o primeiro avanço e até metade dos avanços. Empates de *score* foram resolvidos pela ordem histórica, critério conservador: dentro de um empate, a ordenação proposta reproduz a praticada.

A hipótese é considerada sustentada quando o intervalo de confiança de 95% da diferença entre as duas taxas, no recorte principal, se situa inteiramente acima de zero; inconclusiva quando a estimativa corrigida favorece o *score*, mas o intervalo inclui o zero; e não sustentada quando a estimativa é igual ou desfavorável ao *score*. A regra é unilateral porque a hipótese é direcional. O intervalo provém das mesmas 2.000 reamostras: em cada uma, a diferença é medida apenas nos leads que ficaram de fora da reamostra, e o intervalo corresponde aos percentis 2,5 e 97,5 dessas medidas; por isso, não é centrado na estimativa corrigida e é conservador. A escolha foi feita com base em um estudo de simulação conduzido antes do acesso à base, no qual os intervalos medidos sem reajuste do modelo ou na própria amostra de ajuste cobriram o valor verdadeiro em menos de 95% das vezes e declararam indevidamente a hipótese sustentada, enquanto a medição fora da reamostra manteve a cobertura acima de 95%, sem declarações indevidas. A mesma simulação mostrou que, com 150 a 250 leads, a regra não alcançava o desfecho sustentado, de modo que se registrou, antes da análise, que a classificação esperada era inconclusiva, em razão do tamanho da amostra, e não do mérito do modelo.

Como análise de apoio, que não decide a hipótese, comparou-se a AUC das duas ordenações, com a ordem histórica tratada como *ranking*, e com intervalo pareado obtido nos mesmos leads deixados de fora de cada reamostra. A comparação retrospectiva pressupõe que o desfecho de cada lead não dependa da posição em que foi abordado; para verificá-lo, a taxa de avanço foi comparada entre a primeira e a segunda metade do período.

A produção nacional sobre o tema foi levantada nas bases Spell, SciELO (coleção Brasil) e Google Acadêmico, em 18 de novembro de 2026, com os termos *lead scoring*, qualificação de leads, prospecção B2B, modelo preditivo conversão vendas e pontuação de leads, entre aspas, no

período de 2019 a 2026, recorte que dá continuidade ao mapeamento de Coda RC e Castro GHC (2019). Foram triados por título e resumo todos os registros do Spell e da SciELO e os 20 primeiros resultados de cada termo no Google Acadêmico, em um total de 125 registros, 116 após a remoção de duplicatas, dos quais 7 foram selecionados para exame do texto completo; um trabalho localizado em busca complementar foi incluído pela proximidade com o objeto do estudo (COSTA OS e GOUVEIA LB, 2021).

A pesquisa utilizou exclusivamente registros produzidos pela operação em sua atividade regular, sem interação com pessoas para fins de pesquisa. Antes da análise, a base foi anonimizada, com a remoção de nomes, contatos, CNPJ, razão social, endereços de sites e anotações livres, e cada lead recebeu um código sequencial. O estudo enquadra-se nas hipóteses do art. 1º, parágrafo único, da Resolução nº 510/2016 do Conselho Nacional de Saúde que dispensam o registro e a avaliação pelo sistema CEP/Conep: a pesquisa voltada ao aprofundamento teórico de situações surgidas na prática profissional, sem revelar dados que identifiquem os sujeitos (inciso VII), e a pesquisa com bancos de dados de informações agregadas, sem identificação individual (inciso V) (BRASIL, 2016). Os dados anonimizados não são considerados dados pessoais para os fins da Lei Geral de Proteção de Dados Pessoais (BRASIL, 2018). O uso dos registros foi autorizado pela empresa, mediante declaração de anuência, e a extração e a anonimização foram feitas com apoio de ferramenta de inteligência artificial generativa, no ambiente em que a operação já processa esses registros.

4 RESULTADOS E DISCUSSÃO

Os registros da operação reuniam 183 organizações mapeadas. Foram excluídas 82 que não chegaram a ser abordadas, 5 abordadas por outro integrante da operação, 2 cujo número de contato se mostrou inválido antes do envio e 5 de outro segmento de clientes, presentes na mesma lista. A amostra analítica totalizou 89 leads, abordados entre 17 de agosto e 23 de outubro de 2026, com desfechos observados até a extração, em 16 de novembro de 2026.

Dos 89 leads, 15 responderam com abertura para conversa, o que corresponde a uma taxa de avanço de 16,9%; 7 (7,9%) recusaram a abordagem e 67 (75,3%) não responderam. Entre os 67 leads sem resposta, 12 receberam apenas resposta automática. Nenhuma reunião foi registrada no período. Duas respostas de classificação ambígua foram enquadradas como recusa pela regra definida nos métodos, por declararem que a organização não precisava da solução ou por tomarem a abordagem por outro assunto (Tabela 1).

Tabela 1 – Caracterização da amostra analítica

Característica	Leads	Avanços	Taxa de avanço (%)
Total	89	15	16,9
Segmento de maior volume, segundo o operador	28	2	7,1
Demais segmentos	61	13	21,3
Um contato recebido	47	12	25,5
Dois contatos recebidos	42	3	7,1
Primeira metade do período	44	10	22,7
Segunda metade do período	45	5	11,1

Fonte: SILVA JLM e LIMA ICC, 2026; dados extraídos dos registros da operação.

Os leads de outros segmentos somaram 61 casos, distribuídos em nove categorias de atuação, e foram agrupados conforme os métodos. Nenhum lead recebeu os três contatos previstos na cadência: 47 receberam apenas o primeiro contato e 42 receberam também um contato de acompanhamento. A taxa de avanço mais alta entre os leads com um único contato decorre do próprio registro, que contabiliza os contatos até a abertura da conversa: quem respondeu ao primeiro contato não recebeu acompanhamento. Por essa razão, o número de contatos foi mantido fora das preditoras. A taxa de avanço foi de 22,7% na primeira metade do período e de 11,1% na segunda, diferença de 11,6 pontos percentuais que não alcança significância estatística (teste exato de Fisher, p = 0,167), mas que é relevante em magnitude e é tratada como limitação.

O memorando e os registros da operação descrevem um processo de priorização organizado em três momentos. No mapeamento, as organizações são localizadas em buscadores de mapas e em redes sociais e admitidas na lista quando o segmento de atuação e o volume aparente de demanda indicam aderência à solução ofertada; as que não apresentam esse perfil são descartadas ainda nessa etapa. Na triagem, uma ferramenta de inteligência artificial generativa, configurada com uma base de conhecimento sobre o negócio, examina o perfil público de cada organização, suas publicações e o público predominante, e devolve o meio de contato, o segmento e um nível de prioridade, em quatro categorias: alta, média, baixa e "verificar mais". O operador revisa esse resultado e o altera quando discorda da avaliação, como no caso registrado de um perfil classificado com prioridade alta pelo tamanho do público, retirado da lista por não ter a necessidade que a solução atende. Na abordagem, os leads são contatados por nível de prioridade, dos mais altos para os mais baixos, e, dentro de cada nível, na ordem da triagem, em ritmo de cerca de 22 contatos por dia.

Quanto aos fatores que favorecem a resposta, o memorando atribui o maior peso à forma da abordagem: iniciar o contato com uma pergunta, oferecer a solução sem insistência, gerar interesse e tratar de um problema próprio do segmento do lead. Entre as características das organizações, o operador associa maior chance de resposta aos segmentos de maior volume de demanda e aos contatos feitos pelo número pessoal do responsável. Em sentido oposto, considera que o número de seguidores e a aparência mais ou menos sofisticada do perfil nas redes sociais não alteram a chance de resposta, embora utilize os seguidores para estimar o volume de demanda na etapa de mapeamento (Quadro 1).

Quadro 1 – Critérios de priorização identificados na etapa qualitativa

Critério	Explícito ou implícito	Conhecido antes do contato	Registrado na base	Uso no estudo
Segmento de atuação	Explícito	Sim	Sim	Preditora, agrupada em duas categorias
Volume aparente de demanda	Implícito	Sim	Não	Representado pelo segmento
Seguidores e público nas redes sociais	Implícito	Sim	Não	Não incluído, por falta de registro
Nível de prioridade da triagem	Explícito	Sim	Não, de forma estruturada	Orienta a ordenação de referência
Contato pelo número pessoal do responsável	Implícito	Em parte	Não	Não incluído, por falta de registro
Atendimento automático no número contatado	Implícito	Não, apenas no contato	Sim, no desfecho	Excluído, por configurar vazamento
Aparência do perfil nas redes sociais	Implícito	Sim	Não	Considerado irrelevante pelo operador
Forma da abordagem	Explícito, nos modelos de mensagem	Não se aplica, por ser comum a todos	Sim	Descrição do processo

Fonte: SILVA JLM e LIMA ICC, 2026.

O Quadro 1 indica que, dos oito critérios identificados, apenas o segmento de atuação reúne as três condições exigidas de uma preditora: é conhecido antes do contato, está registrado de forma estruturada e varia entre os leads. Os demais critérios associados pelo operador à organização, como o volume aparente, os seguidores e o tipo de número, orientam a decisão sem deixar registro. O nível de prioridade, embora explícito, também não foi preservado de forma estruturada, de modo que a ordenação de referência foi reconstituída pela data efetiva do primeiro contato.

O modelo, ajustado com essa única preditora, estimou coeficiente de –1,259 para o segmento de maior volume de demanda, o que corresponde a uma razão de chances de 0,284 (IC 95%: 0,059 a 1,356; p = 0,115): na amostra, a chance de avanço desse segmento foi cerca de 72%

menor que a dos demais. O intervalo, contudo, inclui o valor 1, e o teste da razão de verossimilhança não rejeita, ao nível de 5%, o modelo sem preditoras (χ^2 = 3,12; gl = 1; p = 0,077). O pseudo-R^2 de McFadden, de 0,039, indica baixa capacidade explicativa, e o teste de Hosmer-Lemeshow não aponta falta de ajuste (χ^2 = 3,93; p = 0,687), resultado pouco informativo em um modelo com uma única preditora binária. Não houve separação completa ou quase completa. Com as duas respostas ambíguas classificadas como abertura, a razão de chances passa a 0,615 (IC 95%: 0,181 a 2,091; p = 0,437): a direção da associação se mantém, mas sua magnitude diminui.

O *score* de cada lead corresponde à probabilidade estimada pelo modelo, que assume apenas dois valores: 21,3% para os demais segmentos e 7,1% para o segmento de maior volume. A AUC aparente foi de 0,609. A correção de otimismo, calculada em 2.000 reamostras, foi de 0,006, o que resulta em AUC corrigida de 0,603, abaixo da faixa de 0,7 a 0,8 que Hosmer DW, et al. (2013) associam a discriminação aceitável. Nenhuma reamostra foi descartada, e 1.997 delas deixaram fora leads suficientes para a medição fora da reamostra. A Tabela 2 confronta a ordenação por *score* e a ordem histórica de abordagem, e a Figura 1 apresenta as curvas de ganho das duas ordenações.

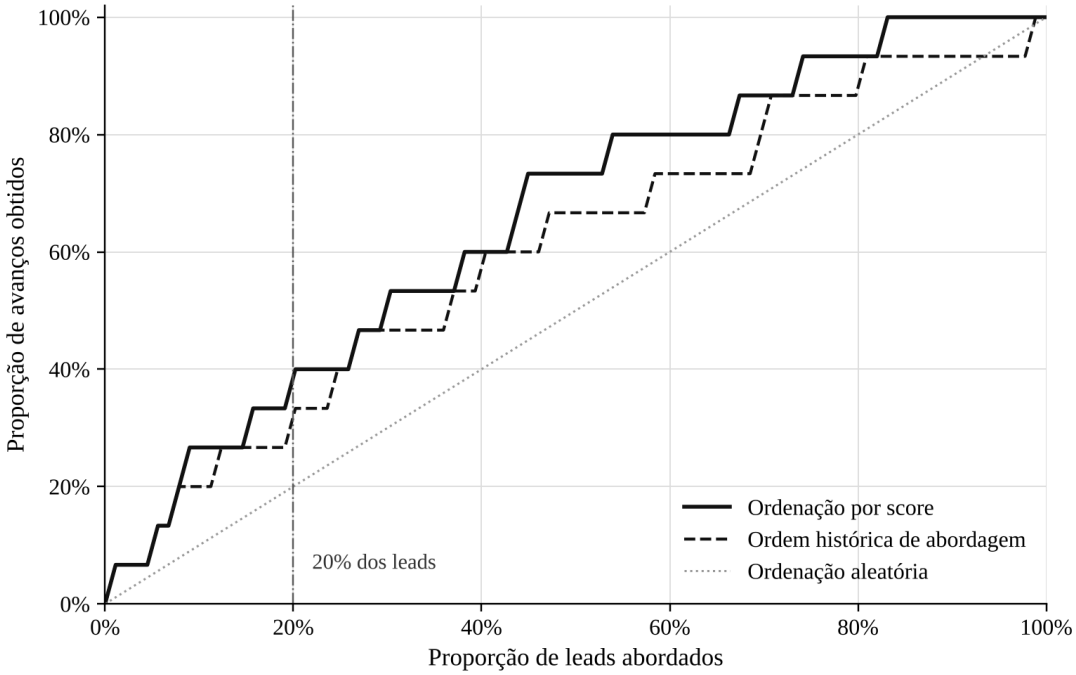
Tabela 2 – Desempenho da ordenação por score e da ordem histórica de abordagem

Métrica	Ordenação por score	Ordem histórica	Diferença
Taxa de avanço nos primeiros 10% (9 leads)	44,4%	33,3%	+11,1 p.p.
Taxa de avanço nos primeiros 20% (18 leads), métrica principal	33,3%	27,8%	+5,6 p.p.; corrigida: +6,3 p.p. (IC 95%: −16,7 a +16,7)
Taxa de avanço nos primeiros 30% (27 leads)	29,6%	25,9%	+3,7 p.p.
Leads abordados até o primeiro avanço	1	1	0
Leads abordados até metade dos avanços	27	33	−6
AUC	0,603 (corrigida)	0,625	−0,027 (IC 95%: −0,333 a +0,280)

Fonte: SILVA JLM e LIMA ICC, 2026; dados extraídos dos registros da operação.

Nota: as diferenças nos recortes de 10% e 30% são aparentes e constituem análise de sensibilidade. Valores arredondados a uma casa decimal; a diferença corrigida foi calculada com as estimativas sem arredondamento (aparente de +5,56 p.p. e otimismo de −0,76 p.p.).

Figura 1 – Curva de ganho da ordenação por score e da ordem histórica de abordagem



Fonte: SILVA JLM e LIMA ICC, 2026; dados extraídos dos registros da operação.

No recorte principal, correspondente aos primeiros 20% da fila, a ordenação por *score* registrou taxa de avanço de 33,3%, ante 27,8% da ordem histórica, diferença aparente de 5,6 pontos percentuais. A correção de otimismo foi negativa, de –0,8 ponto percentual, e a diferença corrigida foi de 6,3 pontos percentuais, valor obtido com as estimativas sem arredondamento (5,56 + 0,76 = 6,32), com intervalo de confiança de 95%, medido fora da reamostra, de –16,7 a +16,7 pontos percentuais. Como a estimativa favorece a ordenação por *score* e o intervalo inclui o zero, a hipótese é classificada como inconclusiva, segundo a regra fixada nos métodos. Nos recortes de sensibilidade, as diferenças aparentes foram de +11,1 pontos percentuais nos primeiros 10% e de +3,7 pontos percentuais nos primeiros 30%.

Quanto ao esforço, as duas ordenações obtiveram o primeiro avanço já no primeiro lead abordado. Para alcançar metade dos avanços, isto é, 8 dos 15, a ordenação por *score* exigiria a abordagem de 27 leads, e a ordem histórica exigiu 33, uma redução de 6 leads, ou 18% do esforço até esse ponto. Na Figura 1, a curva do *score* não fica abaixo da curva da ordem histórica em nenhum ponto do percurso; o maior afastamento, de 13,3 pontos percentuais, ocorre entre 45% e 69% dos leads abordados.

Na análise de apoio, a ordem histórica, tratada como *ranking*, apresentou AUC de 0,625, valor próximo ao do *score* corrigido, de 0,603. A diferença pareada, de –0,027, tem intervalo de

confiança de −0,333 a +0,280, o que não indica diferença de discriminação entre as duas ordenações. Na análise de sensibilidade com as duas respostas ambíguas classificadas como abertura, a diferença corrigida no recorte principal foi de +3,9 pontos percentuais (IC 95%: −42,9 a +16,7), e o veredito permaneceu inconclusivo.

O estudo buscou verificar se um *score* estimado por regressão logística melhoraria a priorização praticada pela operação. A estimativa favoreceu o *score*, com diferença corrigida de 6,3 pontos percentuais no recorte principal e redução de 6 leads no esforço até metade dos avanços, mas o intervalo de confiança incluiu o zero, e a hipótese foi classificada como inconclusiva, desfecho declarado como esperado antes da análise. Com 15 eventos, o critério de Peduzzi P, et al. (1996) limitou o modelo a uma única preditora, e a ordenação por *score* se reduziu a separar dois grupos de leads, o que limita, desde o início, o ganho que o modelo poderia oferecer. A flexibilização admitida por Vittinghoff E e McCulloch CE (2007) não foi empregada, porque não havia outras informações registradas antes do contato que justificassem parâmetros adicionais.

O resultado que contrariou a expectativa foi a direção da associação entre o segmento e o avanço. O memorando, redigido antes do ajuste do modelo, apontava o segmento de maior volume de demanda como o de melhor encaixe, e esse segmento apresentou a menor taxa de avanço da amostra, 7,1% contra 21,3% dos demais. A pesquisa não é explicativa, e a associação não é estatisticamente significativa, de modo que o achado não autoriza conclusões causais. Ele indica, contudo, que o critério que orienta o mapeamento, a necessidade aparente da solução, não coincide necessariamente com a disposição para responder à abordagem: parte das recusas registradas declarou já dispor de equipe ou de sistema para a função que a solução atende.

Esse descompasso entre o critério tácito e o desfecho observado é o tipo de situação que a literatura usa para sustentar a vantagem dos modelos preditivos sobre o julgamento (WU M, et al., 2024). No estudo de González-Flores L, et al. (2025), o modelo superou o *scoring* manual praticado pela empresa, com base mais ampla e com variáveis de origem e de status do lead, indisponíveis aqui. Neste estudo, a vantagem não pôde ser demonstrada, e a AUC corrigida situou-se abaixo da faixa considerada aceitável. A diferença é coerente com a perspectiva de D'Haen J e Van den Poel D (2013), para quem a predição na prospecção B2B depende do retorno acumulado sobre os leads abordados: a operação estudada encontra-se na fase inicial desse processo, e o modelo deve ser tratado como versão a ser recalibrada à medida que o histórico cresce. A interpretabilidade defendida por Rudin C (2019) mostrou utilidade prática: por ser

legível, o modelo expôs de imediato que sua única regra contraria a crença do operador, informação que um modelo opaco não tornaria visível.

A ordem histórica, tratada como *ranking*, apresentou AUC de 0,625, o que sugere que a priorização praticada também carrega alguma informação sobre o desfecho. O resultado admite duas leituras que os dados não permitem separar. A primeira é que a triagem assistida por inteligência artificial, com a revisão do operador, de fato antecipou parte dos avanços. A segunda é que a queda da taxa de avanço ao longo do período, de 22,7% para 11,1%, favorece qualquer ordenação que coloque primeiro os leads abordados mais cedo. Essa queda pode decorrer do esgotamento dos leads de maior prioridade, abordados antes, de variações sazonais ou de mudanças nas mensagens, e compromete em parte o pressuposto da comparação retrospectiva.

A etapa qualitativa produziu um resultado de natureza diferente. Dos oito critérios que o operador emprega, apenas o segmento de atuação ficou registrado de forma estruturada; o volume aparente, os seguidores, o tipo de número e o próprio nível de prioridade orientam a decisão sem deixar registro. A limitação do modelo decorre, assim, menos da técnica do que da informação disponível, o que tem implicação prática direta para operações de pequeno porte: a priorização preditiva depende de registrar, no momento da triagem, os atributos que o operador já observa. O registro do número de contatos ilustra o cuidado inverso: os leads que receberam um único contato avançaram mais porque quem respondeu ao primeiro contato não recebeu acompanhamento, e incluir essa variável seria um caso de vazamento, nos termos de Kaufman S, et al. (2012).

No contexto brasileiro, em que a inteligência artificial em relações B2B tem sido tratada sobretudo em revisões de literatura (BARROS PM e SOUZA CA, 2025) e a análise de dados aparece pouco na produção sobre marketing B2B (CODA RC e CASTRO GHC, 2019), o estudo oferece uma aplicação empírica com dados de uma operação real, ainda que de pequena escala. Entre suas forças estão as regras de decisão fixadas antes da análise, a correção de otimismo, o intervalo medido fora da reamostra e o memorando registrado antes do ajuste do modelo, salvaguarda contra o viés de quem pesquisa a própria operação (BRANNICK T e COGHLAN D, 2007). Entre suas fraquezas estão a amostra de 89 leads com 15 eventos, a preditora única, a interrupção das cadências antes do terceiro contato, a queda da taxa de avanço ao longo do período, a ausência de registro estruturado da prioridade, a restrição da amostra aos leads efetivamente abordados e o caráter de estudo de caso único, que não autoriza generalização.

Para a operação, o resultado não justifica substituir a priorização praticada pela ordenação por *score*, mas indica três usos imediatos. O primeiro é registrar na triagem os atributos que hoje orientam a decisão sem deixar rastro, a começar pelo nível de prioridade e pelo tipo de número de contato. O segundo é reestimar o modelo à medida que os avanços se acumulem, até que o número de eventos permita incluir mais de uma preditora. O terceiro é rever o pressuposto de que o segmento de maior volume é o de melhor encaixe, que os dados desta amostra não confirmaram.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa partiu do problema de como um modelo de *lead scoring* baseado em regressão logística pode apoiar a priorização de leads em um processo de prospecção B2B. Na operação estudada, o modelo pode apoiar a priorização, mas a evidência obtida não permite afirmar que a ordenação por *score* supere a priorização praticada. A ordenação por *score* registrou taxa de avanço maior no recorte principal e exigiria menos leads abordados para alcançar metade dos avanços, mas a incerteza da estimativa não exclui a ausência de ganho, e a hipótese foi classificada como inconclusiva, desfecho que a própria metodologia havia indicado como o mais provável, dado o tamanho da amostra.

A principal contribuição do estudo está menos no modelo obtido do que no que ele revelou sobre a operação. O segmento que o operador considerava de melhor encaixe apresentou a menor taxa de avanço, e a maior parte dos critérios efetivamente empregados na priorização não deixa registro, o que impede que sejam avaliados. Para operações de pequeno porte, a implicação é que a priorização preditiva começa pelo registro sistemático, no momento da triagem, das informações que o operador já observa. O estudo contribui também com um procedimento reproduzível, com regras de decisão definidas antes da análise, validação interna por correção de otimismo e código disponível em repositório público, aplicável a bases maiores.

As limitações decorrem sobretudo do tamanho da amostra, que restringiu o modelo a uma preditora e ampliou a incerteza das estimativas, e do caráter de estudo de caso único, que não permite generalizar os resultados. Como trabalhos futuros, sugere-se que a operação registre na triagem o nível de prioridade, o tipo de número de contato e os indicadores de audiência; que o modelo seja reestimado à medida que o histórico cresça e validado de forma prospectiva, com a alternância entre a ordenação por *score* e a priorização praticada em períodos comparáveis; e

que a classificação produzida pela triagem assistida por inteligência artificial seja avaliada como ordenação própria, confrontada com o modelo estatístico sobre os mesmos leads.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Na elaboração deste trabalho, foi utilizada a ferramenta de inteligência artificial generativa Claude, desenvolvida pela Anthropic, no modelo Claude Opus 5.5, entre agosto e novembro de 2026, na estruturação do texto, na redação de versões preliminares revisadas pelo primeiro autor, na verificação das referências, na triagem da busca bibliográfica, no desenvolvimento e na execução do código de análise e do estudo de simulação, na elaboração de tabelas e figuras e na adaptação do texto às normas da revista. As decisões sobre problema, método, variáveis, classificação dos desfechos e critérios de decisão foram tomadas pelo primeiro autor. A ferramenta processou os registros da operação apenas para a extração e a anonimização da base, e a análise utilizou exclusivamente a base anonimizada. Os autores assumem integral responsabilidade pelo conteúdo.

REFERÊNCIAS

BARROS PM, SOUZA CA. A contribuição da inteligência artificial nos relacionamentos B2B entre empresas de base tecnológica: uma revisão sistemática de literatura. In: Seminários em Administração (SEMEAD), 28., São Paulo, 2025. Anais... São Paulo: FEA-USP, 2025.

BRANNICK T, COGHLAN D. In defense of being "native": the case for insider academic research. Organizational Research Methods, 2007; 10(1): 59-74.

BRASIL. Conselho Nacional de Saúde. Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais. Diário Oficial da União, Brasília, DF, 24 maio 2016; Seção 1: 44-46.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018.

BRUGALLI GP, et al. A link prediction-based method towards lead management. In: Symposium on Knowledge Discovery, Mining and Learning (KDMiLe), 10., Campinas, 2022. Anais... Porto Alegre: Sociedade Brasileira de Computação, 2022; 122-129.

CODA RC, CASTRO GHC. Marketing business-to-business: análise da produção científica brasileira de 2008 a 2018. Revista de Administração de Empresas, 2019; 59(4): 258-270.

COSTA OS, GOUVEIA LB. Prevendo a taxa de conversão de leads no setor da educação com técnicas de aprendizagem de máquina. RECIMA21: Revista Científica Multidisciplinar, 2021; 2(9): e29712.

CRESWELL JW, CRESWELL JD. Projeto de pesquisa: métodos qualitativo, quantitativo e misto. 5. ed. Porto Alegre: Penso, 2021.

D'HAEN J, VAN DEN POEL D. Model-supported business-to-business prospect prediction based on an iterative customer acquisition framework. Industrial Marketing Management, 2013; 42(4): 544-551.

GIL AC. Como elaborar projetos de pesquisa. 7. ed. São Paulo: Atlas, 2022.

GONZÁLEZ-FLORES L, et al. The relevance of lead prioritization: a B2B lead scoring model based on machine learning. Frontiers in Artificial Intelligence, 2025; 8: 1554325.

HOSMER DW, et al. Applied logistic regression. 3. ed. Hoboken: Wiley, 2013.

KAUFMAN S, et al. Leakage in data mining: formulation, detection, and avoidance. ACM Transactions on Knowledge Discovery from Data, 2012; 6(4): 15.

LIMA AFO, et al. Qualificação de potenciais clientes de uma emissora de televisão utilizando a lógica fuzzy. Mosaico: Revista Multidisciplinar de Humanidades, 2024; 15(1): 181-198.

MONAT JP. Industrial sales lead conversion modeling. Marketing Intelligence & Planning, 2011; 29(2): 178-194.

OLIVEIRA AVM, SILVA PHL. Implementação de modelo de qualificação de leads de um funil B2B utilizando técnicas de mineração de dados. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Universidade Federal Rural do Semi-Árido, Mossoró, 2022.

PEDREGOSA F, et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 2011; 12: 2825-2830.

PEDUZZI P, et al. A simulation study of the number of events per variable in logistic regression analysis. Journal of Clinical Epidemiology, 1996; 49(12): 1373-1379.

RABELO AR. Otimização do processo de vendas com a aplicação de uma RNA para detecção de oportunidades abertas de vendas para empresa Soprojetos®. Dissertação (Mestrado Profissional em Engenharia de Processos) – Instituto de Tecnologia. Universidade Federal do Pará, Belém, 2019.

REIS VG. Agente inteligente para qualificação de leads no setor automotivo utilizando machine learning. Trabalho de Conclusão de Curso (Graduação em Ciência da Computação) – Faculdade de Computação. Universidade Federal de Uberlândia, Uberlândia, 2025; 75 p.

RUDIN C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 2019; 1(5): 206-215.

SABNIS G, et al. The sales lead black hole: on sales reps' follow-up of marketing leads. Journal of Marketing, 2013; 77(1): 52-67.

SEABOLD S, PERKTOLD J. Statsmodels: econometric and statistical modeling with Python. In: Python in Science Conference, 9., Austin, 2010. Proceedings... SciPy, 2010; 92-96.

SEBRAE. Pequenos negócios somam 24 milhões de empresas ativas. Agência Sebrae de Notícias, 26 jun. 2026. Disponível em: https://agenciasebrae.com.br/dados/pequenos-negocios-somam-24-milhoes-de-empresas-ativas/. Acesso em: 21 out. 2026.

STEYERBERG EW, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology, 2001; 54(8): 774-781.

VITTINGHOFF E, McCULLOCH CE. Relaxing the rule of ten events per variable in logistic and Cox regression. American Journal of Epidemiology, 2007; 165(6): 710-718.

WU M, et al. The state of lead scoring models and their impact on sales performance. Information Technology and Management, 2024; 25(1): 69-98.