

APLICAÇÕES DE TÉCNICAS ENSEMBLE DE APRENDIZAGEM DE MÁQUINA NA INDÚSTRIA BANCÁRIA

Resumo: Tecnologias de Inteligência Artificial (IA) terão um enorme impacto no setor de serviços financeiros, com diversas implicações, envolvendo desde a redefinição de processos, produtos, serviços até a transformação da forma de interação com clientes. Apesar dos benefícios ainda tímidos da IA, considerando o impacto que já está sendo observado nos mercados financeiros e o potencial poder de transformação disruptiva de toda a indústria bancária, as instituições financeiras em todo o mundo estão realizando investimentos de grande escala em inteligência artificial. Os bancos que obterão mais benefícios das inovações em IA são aqueles que estiverem mais preparados e dispostos a fazerem mudanças e adaptarem suas abordagens com relação a pessoas, processos e dados. No contexto das inovações em finanças, este trabalho tem duas dimensões principais. A primeira dimensão objetiva apresentar uma visão geral de inovações financeiras, trazendo uma perspectiva histórica e discutindo elementos de teoria de finanças. Assim, as inovações financeiras surgem em um contexto em que premissas como, por exemplo, ausência de custos de transação ou existência de mercados completos, são violadas. A segunda dimensão desse trabalho busca ilustrar o uso de técnicas de IA, mais particularmente de aprendizagem de máquina em aplicações voltadas ao segmento bancário. São aplicados métodos *ensemble* baseados em *bagging*, *boosting* e *random forest* para problemas de classificação em duas bases de dados: (i) de tomadores inadimplentes e não inadimplentes em operações de crédito pessoal, (ii) de clientes que aderem ou não aderem a esforços de *marketing* para realização de depósitos bancários. Os resultados da análise empírica sugerem que não há um método de classificação que seja comumente superior, indicando que técnicas de IA devem ser adaptadas aos problemas específicos, considerando ainda as características da base de dados disponível.

Palavras-chave: Inovação em finanças, Inteligência artificial, Aprendizagem de máquina, Métodos *ensemble*, Aplicações em bancos.

Abstract: Artificial Intelligence (AI) technologies will have a massive impact on the financial services sector, with several implications, ranging from the redefinition of processes, products, services to the transformation of the way of interacting with customers. Despite the still timid benefits of AI, considering the impact already being seen in the financial markets and the potential disruptive, transformational power of the entire banking industry, financial institutions worldwide are making large-scale investments in artificial intelligence. The banks that will benefit from AI innovations are those that are most prepared and willing to make changes and adapt their approaches to people, processes, and data. In the context of innovations in finance, this work has two main dimensions. The first dimension aims to present an overview of financial innovations, bringing a historical perspective and discussing elements of finance theory. Thus, financial innovations arise in a context in which assumptions, for example, the absence of transaction costs or the existence of complete markets, are violated. The second dimension of this work seeks to illustrate the use of AI techniques, more particularly machine learning in applications aimed at the banking segment. Ensemble methods based on bagging, boosting, and random forest are applied to classification problems in two databases: (i) defaulting and non-defaulting borrowers in personal loan transactions, (ii) customers who adhere, or do not adhere to marketing efforts of banks. The results of the empirical analysis suggest that there is no classification method that is commonly superior, indicating that AI techniques must be adapted to specific problems, taking into account the characteristics of the available database.

Keywords: Innovation in finance, Artificial intelligence, Machine learning, Ensemble methods, Uses in banking.

1. Introdução

Tecnologias de Inteligência Artificial (IA) terão um enorme impacto no setor de serviços financeiros, com diversas implicações como, por exemplo, (i) na redefinição de processos, produtos, serviços e (ii) na transformação da forma de interação com clientes (Jubraj, Graham e Ryan, 2018). Estudo do *Bank of International Settlements* (BIS) indica que a inteligência artificial traz revoluções substanciais no segmento financeiro, redesenhando modelos e mecanismos de comunicação e interação entre os participantes do mercado (BIS, 2019).

Porém, atualmente, bancos estão obtendo benefícios limitados do uso de IA, dadas as dificuldades em adotar e incorporar efetivamente essa nova tecnologia nos negócios, principalmente em função da necessidade de mudança nos modelos de gestão (KPMG, 2020). De acordo com estudo da KPMG (2020), bancos que superarem essa dificuldade de gestão terão vantagens competitivas contundentes.

A despeito dos benefícios ainda tímidos da IA, considerando o impacto que já está sendo observado nos mercados financeiros e o potencial poder de transformação disruptiva de toda a indústria bancária, as instituições financeiras em todo o mundo estão realizando investimentos de grande escala em inteligência artificial (McWaters e Galaski, 2018).

Neste contexto, Plaschke, Seth e Whiteman (2018) sugerem que a inteligência artificial e, em especial, os algoritmos de aprendizagem de máquina, transformarão a função financeira, exigindo um posicionamento estratégico para competir em um mercado que pode criar valor substancial nos próximos anos. Mais particularmente, conforme estimativas da *McKinsey*

Global Institute, ferramentas de inteligência artificial podem gerar valor adicional de mais de USD 250 bilhões na indústria bancária (Babel, Buehler, Pivonka, Richardson e Waldron, 2019).

De acordo com Jubraj, Graham e Eve (2018), corroborando os resultados da análise da KPMG (2020), os bancos que obterão mais benefícios das inovações em IA são aqueles que estiverem mais preparados e dispostos a fazerem mudanças e adaptarem suas abordagens com relação a pessoas, processos e dados. No contexto de adoção de IA, conforme estudo elaborado por McCauley (2020), as empresas de serviços financeiros estão entre as pioneiras, logo depois das empresas de tecnologia de informação e telecomunicação e antes das empresas de bens de consumo e varejo.

É importante destacar que os ganhos de valor advindos da inovação financeira podem ser acompanhados também por um aumento dos riscos potenciais ao mercado. Por exemplo, conforme sugerem Muniesa e Lenglet (2013), a inovação financeira teve papel proeminente na crise de crédito do final da década passada, fomentando o debate sobre as responsabilidades da indústria de serviços bancários.

As inovações no segmento financeiro como, por exemplo, derivativos de crédito e *trading* algorítmico refletem não somente o uso de criatividade para o desenvolvimento de soluções financeiras e exploração de oportunidades de mercado, como também externalidades negativas, falhas regulatórias e riscos tecnológicos (Millo e MacKenzie, 2009; Muniesa e Lenglet, 2013; Taleb, 2007).

As inovações financeiras no mercado de securitização como, por exemplo, *Collateralized Debt Obligations* (CDOs) e *Credit Default Swaps* (CDS), aumentaram a tolerância ao risco de

investidores em títulos ligados a hipotecas (Immergluck, 2009) que, atrelada a uma regulação mais leniente, criou condições para a crise econômica global de 2007. Apesar da velocidade cada vez maior das inovações, o segmento financeiro não desenvolveu uma cultura de escrutínio público e precaução técnica como outros segmentos como, por exemplo, biomedicina, biotecnologia, telecomunicações, etc (Callon e Lacoste, 2011; Muniesa e Lenglet, 2013).

Neste sentido, o uso de IA em aplicações no segmento financeiro deve ser avaliado ao mesmo tempo com motivação e cautela. Os ganhos de produtividade e de poder analítico que a IA propicia aos gestores financeiros podem vir acompanhados de armadilhas que, por sua vez, podem se refletir em riscos sistêmicos com potencial de impacto considerável na economia global.

Desta forma, deve-se evitar o uso indiscriminado de algoritmos de IA, focando-se no entendimento das premissas teóricas e das implicações práticas não somente do ponto de vista gerencial como também social. Por exemplo, as inovações financeiras baseadas em IA devem ser acompanhadas de debate sobre privacidade e ética, uma vez que grande parte da matéria prima para modelos de aprendizagem de máquina envolve dados, muitas vezes sigilosos e confidenciais, de indivíduos e empresas.

No contexto das inovações em finanças, este trabalho tem duas dimensões principais. A primeira dimensão objetiva apresentar uma visão geral de inovações financeiras, trazendo uma perspectiva histórica e discutindo elementos de teoria de finanças. Assim, as inovações financeiras surgem em um contexto em que premissas como, por exemplo, ausência de custos de transação ou existência de mercados completos, são violadas.

A segunda dimensão desse trabalho busca ilustrar o uso de técnicas de IA, mais particularmente de aprendizagem de máquina em aplicações voltadas ao segmento bancário. São aplicados métodos *ensemble* baseados em *bagging*, *boosting* e *random forests* para problemas de classificação de (i) tomadores inadimplentes e não inadimplentes em operações de crédito pessoal, (ii) adesão ou não adesão a esforços de *marketing* para realização de depósitos bancários.

O trabalho está estruturado da seguinte forma. Na próxima seção, é realizada uma discussão teórica sobre inovação financeira, considerando contexto, tipologia e são apresentadas algumas inovações no segmento bancário. Posteriormente, são brevemente descritos os métodos *ensemble* que serão explorados nesse estudo, visando ilustrar o uso de técnicas de inteligência artificial na indústria bancária. Os resultados do estudo são discutidos e finalmente são apresentados os comentários finais, incluindo limitações e sugestões para estudos futuros.

2. Discussão teórica

a. Contexto das inovações em finanças

Conforme Franklin (2012), uma explicação tradicional para a existência de inovação financeira está associada ao potencial preenchimento de lacunas, buscando a correção de algum tipo de ineficiência ou imperfeição de mercado.

Em um mercado perfeito, sem custos de transação, sem custos de informação e sem regulamentação, tanto a intermediação financeira quanto a inovação financeira não existiriam (Blake, 1996). Nestas condições, considerando ainda a inexistência de problemas de razão

moral (Tufano, 2003), os indivíduos poderiam, por si só, replicar, sem custos, todas as atividades dos intermediários financeiros ou todas as características de produtos inovadores (Blake, 1996). Desta forma, os esforços para inovação não criariam valor aos seus desenvolvedores.

Esse argumento de irrelevância da inovação financeira, dentro da perspectiva de os indivíduos ou empresas poderem replicar produtos do mercado, é análogo ao da irrelevância de estrutura de capital, explorada na teoria de finanças pelos trabalhos seminais de Modigliani e Miller (1958, 1963), que trazem à discussão a irrelevância do conceito de propriedade da empresa (Fama, 1980).

Assim, sem imperfeições de mercado, os indivíduos poderiam replicar estratégias de financiamento e de investimento em títulos financeiros, de forma competitiva aos produtos ou serviços ofertados por intermediários financeiros. No entanto, quando mercados são incompletos, as inovações financeiras podem explorar oportunidades de compartilhamento de riscos. Inovações, notadamente por meio da formulação de títulos financeiros, podem também alinhar interesse dos diversos *stakeholders* quando há conflitos de interesse (Franklin, 1992).

Além disso, Merton (1992) discute que uma teoria amplamente aceita sobre os direcionadores da inovação financeira envolve a redução de custos e a transposição de restrições regulatórias tanto em termos de regras tributárias quanto em termos de convenções contábeis. Essa discussão é semelhante à de Franklin (2012) que sugere que as inovações podem ser motivadas por possibilidade de diminuir impostos ou evitar impactos da regulação.

Dadas as imperfeições de mercado, as inovações financeiras têm propiciado diversas contribuições para o mercado. Por exemplo, as inovações têm reduzido os custos de transação de diversas operações financeiras, conferindo maior liquidez (Miller, 1992). Além disso, Allen (2012) exemplifica que inovações financeiras na forma de *venture capital* e *private equity* são também importantes mecanismos com efeitos benéficos, principalmente por propiciar caixa para empreendimentos em estágios iniciais de desenvolvimento, quando estão expostos a elevados níveis de risco.

Muitas empresas americanas, no início de suas operações, se beneficiaram de capital de risco como, por exemplo, Apple, Cisco Systems, Google e Microsoft (Allen, 2012). Assim, sem essas inovações, dificilmente empreendimentos com alto grau de risco conseguiriam financiamento para suas operações.

Allen (2012) também estabelece que as inovações financeiras trouxeram contribuição relevante para empreendimentos sem fins lucrativos como, por exemplo, os voltados à preservação ambiental. Assim, o mercado de carbono pode ser considerado uma inovação financeira associado a questões de preservação ambiental e preocupações com as mudanças climáticas globais.

Outra inovação financeira com impacto ambiental são as STFs (*State Revolving Funds*). Nesse fundo, o governo federal concede um recurso a estados, com 20% de contrapartida dos governos estaduais e com suplementos financeiros de filantropos. Esses recursos são emprestados pelo estado a cidades e organizações que, por sua vez, pagam a dívida por meio de receitas dos projetos ligados ao meio ambiente e de impostos locais (Allen, 2012).

Nesta configuração inovadora, recursos financeiros do governo federal permitem a viabilidade de curto prazo dos projetos que, ao gerarem resultados, tornam-se sustentáveis a médio e longo prazos. Em um fundo específico, *Clean Water State Revolving Fund (CWSRF)*, USD 19,8 bilhões de empréstimos para proteção e restauração de águas, geraram USD 7 bilhões de economia, beneficiando 115 milhões de pessoas e criando 716 mil empregos entre 1987 e 2005 (EPA, 2008).

Um outro exemplo de inovação financeira contribuindo para a esfera de iniciativas sem fins lucrativos envolve o financiamento de pesquisas para cura de doenças que assolam primordialmente países em desenvolvimento (Allen, 2012). Tendo em vista que algumas doenças já erradicadas ou pouco comuns em países desenvolvidos não geram interesse de mercado, formas alternativas de financiamentos a partir de recursos privados representam inovações com impacto social relevante.

De forma mais abrangente, as inovações em produtos e serviços financeiros podem contribuir para a melhoria de desempenho econômico, (i) suprimindo demandas de agentes superavitários e deficitários por produtos que aumentem as oportunidades de compartilhamento ou agrupamento de risco, *hedging* de exposições, transferência espacial ou intertemporal de recursos, (ii) diminuindo custos de transação e aumentando liquidez, (iii) reduzindo problemas de agência, advindos de assimetria informacional e de impossibilidade de monitoramento do comportamento do agente (Miller, 1992).

A literatura tem sugerido que, apesar de a inovação financeira comumente trazer benefícios, pode também gerar problemas (Allen, 2012). Ao mesmo tempo que a redução de custos de transação advinda das inovações financeiras permite um maior acesso aos mercados, implica

também em um maior nível de negociação de curto prazo, ocasionando um menor horizonte de tempo de planejamento para empresas e investidores, podendo trazer maior volatilidade (Miller, 1992). Assim, as inovações podem trazer externalidades negativas, exigindo intervenção de governos para corrigir incentivos (Wieneke, 2016) e distorções.

Essa maior instabilidade nos mercados pode ser fonte de crises. Allen (2012) sugere que algumas inovações financeiras são conduzidas com o objetivo de criar complexidade e explorar o cliente, trazendo exemplos como os mecanismos de securitização e hipotecas sub-prime que amplificaram a crise de 2007.

Estudos têm mostrado que há situações nas quais um produto considerado inovador extrai riqueza, com pouca contribuição ao mercado. Por exemplo, SPARQS emitidos pelo Morgan Stanley da década de 2000, apesar de arriscados, geravam um retorno esperado abaixo da taxa de juros livre de riscos, sem apresentar vantagens em relação a propriedades de *hedging*, liquidez ou benefícios fiscais (Allen, 2012; Henderson e Pearson, 2011).

b. Tipologia das inovações em finanças

A inovação em finanças possui diversas perspectivas e dimensões. Sob um ponto de vista macro, a inovação financeira pode ser entendida como o motor que direciona o sistema financeiro para o objetivo de aprimorar o desempenho da economia real (Merton, 1992).

Sob uma perspectiva voltada a operações, Blake (1996) associa a inovação financeira com um processo de desmontagem e reembalagem de características de produtos financeiros visando a criação de novos instrumentos. Neste contexto, a intermediação financeira e a inovação

financeira estão ligadas a mudanças não disruptivas em serviços, sistemas, organizações e reinvidicações sobre fluxos de caixa (Blake, 1996).

Em contrapartida, sob uma perspectiva mais disruptiva, a inovação financeira representa a ação de criação e popularização de novos instrumentos, tecnologias, instituições e mercados financeiros (Tufano, 2003). Para o autor, as inovações em finanças podem ser classificadas em (i) inovações em produtos e (ii) inovações em processos.

Nesta classificação, as inovações em produtos estão, por exemplo, associadas ao desenvolvimento de (i) novos títulos financeiros, (ii) novos instrumentos derivativos ou (iii) novas formas de agrupamentos ou integração de produtos existentes, enquanto as inovações em processo podem ser tipificadas por (i) novos mecanismos de distribuição de ativos, (ii) novas tecnologias de processamento de transações ou (iii) novos modelos de precificação de operações (Tufano, 2003).

De forma análoga, Ross (1989) estabelece que as inovações financeiras podem, apesar de variadas e difusas, sob uma taxonomia mais definitiva, ser classificadas em (i) novos títulos e novos mercados e (ii) novas estratégias de negociação dinâmicas que exploram esses novos instrumentos.

As inovações financeiras que surgiram no século passado tiveram impulso devido a quatro fatores (Miller, 1992): (i) a movimentação para taxa de juros flutuante, (ii) a revolução na tecnologia de informação, principalmente decorrente do avanço dos computadores, (iii) o represamento de avanços no setor nas décadas anteriores em função do crescimento econômico limitado, reflexo dos efeitos da crise de 1929 e (iv) as consequências da Segunda Guerra

Mundial que, dentre os diversos impactos econômicos, ainda implicou restrições às transações internacionais.

A partir da década de 1970, impulsionada pelo processo de desregulamentação financeira advinda do colapso do Acordo de Bretton Woods (Bavoso, 2016; Miller, 1992), a inovação em finanças gerou diversos instrumentos e estratégias com derivativos como, por exemplo, (i) *swaps*, (ii) *swaptions*, (iii) futuros de índices, (iv) opções negociadas em mercados de bolsa, (v) *butterfly spreads*, etc (Aliber, 1987, Miller, 1992; Ross, 1989).

Mecanismos de diversificação de moeda e cláusulas de opcionalidade também foram sendo agregados a títulos de dívida tradicionais, propiciando a negociação de (i) títulos resgatáveis, (ii) títulos *zero-coupon*, (iii) *eurobonds*, (iv) *eurodollars*, (v) títulos de taxas flutuantes, (vi) *junk bonds*. Outras inovações envolvem ainda (i) securitização de recebíveis, (ii) seguro de carteira, (iii) caixa sintético, (iv) *strips*, (v) *straps*, (vi) *primes*, (vii) *scores*, (viii) estratégias de *trading* dinâmico, (ix) sistemas de *trading*, etc (Aliber, 1987, Miller, 1992; Ross, 1989).

Além disso, o desenvolvimento tecnológico permitiu outras inovações em finanças como, por exemplo, (i) negociação de 24 horas, movimentando centros financeiros como Tóquio, Londres e Nova Iorque ao longo do dia, (ii) transferência eletrônica de fundos, (iii) caixas eletrônicos, etc (Merton, 1992). Conforme destaca Aliber (1987), a inovação mudou estruturas, remodelou os participantes e consolidou centros financeiros.

Mais recentemente, uma nova onda de inovações está invadindo a indústria financeira. Por exemplo, o desenvolvimento tecnológico está trazendo diversos outros *players* ao mercado como as *fintechs* que estão aproveitando espaços ao conseguirem reduzir custos de transação e

conquistar clientes (i) usando massivamente os avanços em tecnologia de informação e comunicação e (ii) explorando algoritmos de inteligência artificial em sua base de clientes.

Além das fintechs, outras empresas de tecnologia estão avançando no segmento financeiro. Apple, Google, Alibaba e mais recentemente Amazon e Facebook já possibilitam transações por meio de seus sistemas de pagamento. Adicionalmente, as empresas de tecnologia têm avançado em outras áreas de finanças, com parcerias pouco prováveis no passado. Ilustrando essa movimentação, a Amazon está se aliando ao JPMorgan e a Berkshire Hathaway para estruturar uma nova empresa de seguro de saúde que atenda seus funcionários (Faustino, 2018).

Iniciativas na área de finanças podem representar atualmente uma pequena parcela dos negócios das grandes empresas de tecnologia. Porém, como o número de usuários em suas plataformas é significativo, essas iniciativas podem se transformar em negócios que movimentam recursos vultosos (Shevlin, 2019). Por exemplo, a empresa financeira do Alibaba, Ant Financial Services Group, possui 600 milhões de usuários e a Amazon tem cerca de 300 milhões de clientes que utilizam o Amazon Pay (Shevlin, 2019).

Acompanhando as inovações, novos mecanismos de gestão e novas regulamentações são necessárias para contemplar as mudanças decorrentes do uso de inteligência artificial (Faustino, 2018). Governos e órgãos reguladores têm buscado enfrentar as incertezas e as apreensões que acompanham o avanço de IA em instituições e mercados (McWaters e Galaski, 2018).

Com relação à regulamentação, a Lei Geral de Proteção de Dados representa importante mecanismo para proteção da privacidade. Com relação ao aumento de competitividade no setor

financeiro, o Banco Central regulamentou recentemente a entrada de *fintechs* no segmento de crédito (Reuters, 2018), autorizando a atuação das dessas empresas em empréstimos de recursos próprios e na intermediação entre agentes superavitários e agentes deficitários (Faustino, 2018).

c. Inteligência artificial na indústria bancária

Considerando impactos dos últimos cinquenta anos, pode-se exemplificar alguns exemplos de inovações que tiveram grande repercussão em finanças: algoritmos para reconhecimento ótico de caracteres foram desenvolvidos em 1970, planilhas eletrônicas em 1980 e sistemas de planejamento financeiro e boletagem eletrônica em 1990 (Heric, 2018).

Além das tecnologias de suporte à função financeira, o avanço da teoria e de técnicas propiciaram aplicações mais avançadas como, por exemplo, modelos de alocação de capital, mecanismos alternativos de captação de recursos e ferramentas de gestão de riscos (Avgouleas, 2015, Bavoso, 2016).

Dentro da indústria de serviços financeiros, McCauley, (2020) apresenta resultados de pesquisa conduzida entre janeiro e fevereiro de 2020, com cerca de 110 alto executivos (C-suite, diretores, chefes de IA, chefes de *analytics*), de grandes empresas nas Américas, Europa, Ásia e África.

A pesquisa indica que os principais usos da IA na área financeira concentram-se em (i) detecção de fraude, (ii) processos e análises financeiras, (iii) segurança cibernética, (iv) personalização de produtos e serviços, (v) atendimento ao cliente e (vi) manutenção de ativos (McCauley, 2020).

Algumas evidências anedóticas podem ser apresentadas para ilustrar o uso de IA na indústria bancária. Em 2016, o Citigroup, após ter falhado no teste de stress do Banco Central americano, utilizou técnicas de aprendizagem de máquina para obter aprovação (Heric, 2018). Além disso, a PicPay, uma fintech de pagamentos, pretende aumentar o número de clientes cadastrados e criar algoritmos de IA que permitam conhecer melhor seus usuários e parceiros, para oferecer produtos e serviços customizados (Romani, 2020). Já o Itaú-Unibanco explora IA para a segurança física das agências por meio do monitoramento da circulação de pessoas (Caputo, 2017a), reforçando os mecanismos de combate a potenciais ações criminosas.

Recentemente, o Bradesco lançou a BIA (Bradesco Inteligência Artificial), *chatbot* cognitivo que tira dúvida de clientes e funcionários. Desenvolvido na plataforma Watson da IBM e também usado no Next, banco digital criado pelo Bradesco, a BIA utiliza IA para aperfeiçoar a experiência dos clientes e visa aumentar receitas advindas dos canais digitais (Mari, 2019).

Os *chatbots* tipicamente buscam aprimorar a interação com o usuário por meio de uma experiência automatizada, que explora tecnologias cognitivas, em aplicações voltadas a responder perguntas, dar orientações e auxiliar humanos a desempenhar atividades (Swaroop, 2019). Na área de finanças, os *chatbots* podem ser usados para abertura de contas, questões sobre tarifas, consultas sobre pagamentos, situações sobre pagamentos (Swaroop, 2019). Usualmente, o sistema de *chatbot* processa a linguagem humana, utilizando IA, buscando interagir e responder às demandas do usuário de forma natural (Caputo, 2017b).

No curto prazo, IA e ML vão contribuir para que *fintechs* aumentem a acurácia e personalização de operações de pagamento, empréstimo e seguros (Columbus, 2019). Por exemplo, Prestige Financial Services, financiadora de automóveis de Utah (EUA) utilizou IA para analisar crédito

para pessoas sem histórico ou com pouco histórico, aumentando taxas de aprovação para pessoas menores de 20 anos em três vezes. (Columbus, 2019)

É importante destacar a relevância cada vez maior de IA no segmento de instituições não-financeiras. Pesquisa com cerca de 1000 CEOs, CFOs, executivos financeiros sêniores e júniores mostrou que as funções financeiras já estão usando ferramentas que exploram dados, notadamente usando análises preditivas e IA (Accenture, 2018).

Outra pesquisa, conduzida com executivos do nível C-suite, indicou que 84% dos respondentes acreditam que as empresas devem alavancar o uso de IA para atingir as metas de crescimento (Awalegaonkar, Ketan, Berkey, Robert, Douglass, Greg, Reilly, 2019). O estudo mostra ainda que cerca de 75% dos executivos temem que suas empresas possam ser completamente alijadas do mercado caso não escalem as aplicações de IA nos próximos cinco anos (Awalegaonkar et al., 2019). Essa previsão não parece ser tão improvável, pois de acordo com Jubraj, Graham e Ryan (2018), mais da metade das empresas do Fortune 500 no início dos anos 2000 encerraram suas atividades.

A despeito de a inovação financeira ter, mais recentemente, ganho espaço especialmente com a evolução tecnológica e a introdução de técnicas de IA, seu processo tem existido até mesmo nos períodos em que os sistemas monetários eram ainda rudimentares, principalmente visando a otimização de transações (Bavoso, 2016). Conforme Tufano (2003), a história mostra que a inovação financeira tem sido persistente no ambiente econômico nos últimos séculos.

A despeito da automatização por meio de robôs e de aprendizagem de máquina ser um tópico contemporâneo que gera discussões acaloradas, muitas tecnologias já existentes ainda são bastante sub-utilizadas (Heric, 2018). Em particular, a falta de integração adequada com a estrutura tecnológica existente pode tornar esforços de adoção de novas tecnologias menos assertivas (Hein, 2018).

3. Metodologia

a. Métodos *ensemble*

Considerando que algoritmos de inteligência artificial e, mais especificamente, modelos de aprendizagem de máquina, constituem inovações que estão alterando substancialmente o ambiente competitivo, esse estudo apresenta aplicações de métodos *ensemble* em problemas comumente enfrentados por instituições financeiras.

Métodos *ensemble* constituem um conjunto de técnicas estatísticas e computacionais que combinam múltiplos algoritmos de aprendizagem para obter um desempenho preditivo superior ao de algoritmos individuais (Opitz e Maclin, 1999; Rokach, 2010). Apesar de métodos *ensemble* poderem ser usados para aprendizagem não-supervisionada, tipicamente são aplicadas em aprendizagem supervisionada, na qual o algoritmo desenvolve um modelo baseado em um aprendizado a partir de observações previamente classificadas ou avaliadas em uma amostra de treinamento.

Diversos estudos têm mostrado a superioridade dos métodos *ensemble* em diferentes aplicações de inteligência artificial para reconhecimento de padrões (Kuncheva e Whitaker, 2003; Sollich

e Krogh, 1996). Em finanças, algoritmos *ensemble* têm sido aplicados a diversos problemas como, por exemplo, de previsão de dificuldades financeiras (Liang, Tsai, Lu, Chang, 2020) ou falência corporativa (du Jardin, 2020), de previsão de preços de ativos ou *commodities* (Wang, Niu, Du, Yang, 2020; Aggarwal, Chandrasekaran, Annamalai, 2020), etc.

Neste estudo, são analisados resultados de técnicas *ensemble* de aprendizagem de máquina, notadamente (i) *bagging*, (ii) *boosting* e (iii) *random forest*, em aplicações voltadas a problemas de classificação de observações em classes ou categorias, utilizando base de dados do segmento bancário.

Problemas de classificação estão direcionados à identificação da classe ao qual uma determinada observação é mais provável de pertencer. Por exemplo, em aplicações de *credit scoring*, pode-se usar técnicas de aprendizagem de máquina para identificar ao grupo de inadimplentes ou adimplentes ao qual um determinado tomador tem maior probabilidade de pertencer. Em aplicações de análise de fraude, busca-se identificar se determinada operação financeira é mais provável de ser legítima ou fraudulenta.

Mais especificamente, nesse estudo são analisados dois bancos de dados publicamente disponíveis que abordam: (i) risco de crédito em operações de empréstimo para pessoas físicas e (ii) adesão a programas de *marketing* de incentivo à realização de depósitos em um determinado banco.

Os resultados são comparados com os da regressão logística, que constitui técnica estatística tradicional e comumente usada como *benchmark* para avaliação de desempenho de algoritmos

de classificação. A seguir, são apresentados, de forma resumida, os fundamentos de cada técnica aplicada às bases de dados.

b. Discussão dos algoritmos

- Regressão logística

Uma das técnicas mais tradicionais para classificação, notadamente em variáveis dicotômicas, é a regressão logística. A classificação envolve, portanto, uma variável Y que pode assumir os valores 0 ou 1. Apesar de sua origem na estatística, pode-se considerar também a regressão logística como uma ferramenta de inteligência artificial, uma vez que existe um mecanismo de aprendizagem em função do ajuste de parâmetros a partir de uma amostra de calibração.

Considerando o vetor de k características \mathbf{x} , obtêm-se parâmetros β_i , $i = 0, 1, \dots, k$ de uma regressão linear em uma amostra de treinamento, a partir da função $\ln(p/(1-p)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$, onde p representa a probabilidade de Y ser igual a 1 e $1-p$ representa a probabilidade de Y ser 0.

A classificação em uma amostra de predição envolve substituir as características \mathbf{x} na equação de regressão. No modelo de regressão logística tradicional, Y é classificada como 1, se $p = 1 / (1 + \exp(-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k))) > 0,5$, e Y é classificada como 0, caso contrário.

Destaca-se que resultados de acurácia de classificação de modelos usualmente são comparados com resultados oriundos do uso de regressão logística. Assim, a regressão logística constitui

um *benchmark* para comparação de ganhos de classificação de outros modelos de aprendizagem de máquina;

- *Bagging*

O *bagging* é um método baseado na geração de um preditor agregado médio, obtido a partir combinações múltiplos preditores, utilizando réplicas *bootstrap* do conjunto de aprendizado (Breiman, 1994, 1996).

Para problemas de classificação, define-se \mathcal{L} como o conjunto de aprendizado contendo os dados $\{(y_n, \mathbf{x}_n), n = 1, \dots, N\}$, sendo y os labels associados às diferentes classes ou categorias. Essas classes dependem das diferentes aplicações como, por exemplo, (i) tomador inadimplente ou não inadimplência em análise de *application scoring*, (ii) fraude, erro humano ou problema de sistemas em gestão de risco operacional, (iii) consumidor indiferente, consumidor pouco motivado, consumidor muito motivado por um determinado produto financeiro, em campanhas de *marketing*.

De acordo com Breiman (1994, 1996). no *bagging*, utiliza-se a função $\varphi(\mathbf{x}, \mathcal{L})$ como um modelo de predição de y considerando-se as características \mathbf{x} . Utilizando-se diferentes amostras *bootstrap* do conjunto de treinamento, obtém-se uma sequência $\{\mathcal{L}_k\}$ que permite a identificação de k preditores $\varphi(\mathbf{x}, \mathcal{L}_k)$. Sendo $\varphi(\mathbf{x}, \mathcal{L})$ utilizado para prever uma classe $j \in \{1, \dots, J\}$, então, considerando o critério de maioria de votos, a classificação envolve uma função agregada $\varphi_A(\mathbf{x}) = \operatorname{argmax}_j N_j$, onde N_j representa o número de amostras em que a função $\varphi(\mathbf{x}, \mathcal{L}_k)$ conduziu a uma classificação na classe j .

Ou seja, o algoritmo de *bagging* estabelece que, a partir de uma dada amostra de aprendizado, são geradas amostras aleatórias com reposição, que passam por uma função de predição da categoria j à qual uma dada observação com características x pertence. Destaca-se que em cada amostra, a predição da observação pode ser diferente e, portanto, a categorização final da observação é dada pela classe em que foi classificada na maioria dos modelos de predição.

Técnicas de *bagging* são utilizadas em finanças em diversas aplicações como, por exemplo, em Yin (2020) que estabelece previsões de *equity premium* e de decisão de carteira ótima e em Jordan, Vivian e Wohar (2017) que estima retornos de mercado.

- *Boosting*

Seguindo a discussão de Freund e Schapire (1995, 1999), no *boosting*, há um conjunto de treinamento $\{(x_i, y_i)\}, i = 1, \dots, m$, no qual cada x_i pertence a um espaço X que denota as características que são usadas para classificação de uma determinada observação i no *label* y_i pertencente a um conjunto Y . Tendo em vista que neste estudo são analisadas situações que envolvem apenas duas classes, por exemplo, inadimplente e não inadimplente, define-se que as possíveis categorias são $Y = \{-1; +1\}$, dada a conveniência dessa notação para a construção do algoritmo.

De modo de semelhante ao *bagging*, no *boosting* uma função de classificação, chamada de algoritmo base $h_t: X \rightarrow \{-1; +1\}$, é aplicada repetidamente em uma série de amostras em diferentes rodadas $t = 1, \dots, T$. Porém, diferentemente do *bagging* onde se utilizam amostras *bootstrap* que independem das amostras anteriores, no *boosting* as novas amostras dependem dos erros de classificação ϵ_t obtidos previamente.

Assim, as novas amostras definidas em cada rodada t têm uma distribuição D_t que estabelece maior probabilidade de incluir observações que foram incorretamente classificadas em rodadas anteriores. Assim, esse mecanismo de alteração da distribuição de probabilidades permite um aprendizado que ajusta a classificação final em função dos resultados anteriores do algoritmo.

Mais especificamente, no algoritmo do AdaBoost, representativo de uma versão do *boosting* proposto por Freund e Schapire (1995, 1999), o erro de classificação é obtido por meio de $\epsilon_t = P_{i \sim D_t}(h_t(x_i) \neq y_i) = \sum_{i=h_t(x_i) \neq y_i} D_t(i)$. Essa medida de erro subsidia a obtenção de um parâmetro de ajuste ou ponderação na classificação, dado por $\alpha_t = \frac{1}{2} \ln \left(\frac{1-\epsilon_t}{\epsilon_t} \right)$.

A nova distribuição de probabilidades, que será a base da amostra a ser analisada na próxima rodada, é dada por $D_{t+1}(i) = D_t(i) \exp(-\alpha_t y_i h_t(x_i)) / Z_t$, onde Z_t é um fator de normalização. Após várias rodadas, a classificação final $H(x)$ gerada pelo algoritmo AdaBoost é obtida aplicando-se a função base h_t em x , ponderando-se os resultados pelos coeficientes α_t de cada rodada de classificação, conforme $H(x) = \text{sign} \sum_{t=1}^T \alpha_t h_t(x)$.

Destaca-se que versões adaptadas do *boosting* são utilizadas em diversos estudos na área de finanças como, por exemplo, na identificação de arbitragem estatística (Krauss, Do e Huck, 2017) e na previsão de falências (Zieba, Tomczak e Tomczak. 2016).

- *Random forest*

As *random forests* constituem uma combinação de preditores baseados em árvores de decisão, na qual cada árvore depende de valores de um vetor aleatório gerado de forma independente e identicamente distribuídas em relação a todas as árvores da floresta (Breiman, 2001).

Seguindo Ho (1995), dadas t árvores, para um ponto específico x , considera-se um nó final $v_j(x)$ da árvore T_j , $j = 1, 2, \dots, t$ ao qual a observação foi classificada. A partir do teorema de Bayes, pode-se estabelecer a probabilidade a posteriori de que x pertence à classe c das n possíveis classes, como $P(c|v_j(x)) = P(c, v_j(x)) / \sum_{l=1}^n P(c_l, v_j(x))$.

Essa probabilidade pode ser obtida por meio de $\hat{P}(c|v_j(x))$, calculada a partir da fração de pontos classificados em c em relação a todos os pontos que são atribuídos a $v_j(x)$. Assim, a função que discrimina os pontos é dada por $g_c(x) = 1/t \cdot \sum_{j=1}^t \hat{P}(c|v_j(x))$ (Ho, 1995).

Aplicações recentes de *random forests* para o segmento de instituições financeiras incluem, por exemplo, (i) campanhas de marketing (Ladyzynski, Zbikowski e Gawrysiak, 2019), (ii) previsão de crises bancárias (Beutel, List e Schweinitz, 2019), (iii) detecção automática de relacionamento entre operações bancárias (González-Carrasco et. al., 2019), detecção de trojans que buscam acesso a sistemas de pagamento e serviços bancários online (Gezer et. al., 2019),.

4. Análise empírica

Nesse trabalho, os métodos *ensemble* são aplicadas a duas bases de dados. A primeira base é composta inicialmente por 32.581 observações simuladas de dados de *bureau* de crédito, e é disponibilizada por Tse (2020). A variável a ser predita envolve a classificação de um cliente pessoa física tornar-se ou não inadimplente ao contrair um empréstimo bancário.

Assim, o objetivo da aplicação é classificar o tomador, com base em características do cliente e da operação, no grupo de inadimplentes ou adimplentes. Na carteira disponibilizada, há uma taxa de inadimplência de cerca de 22% e, portanto, a base é suavemente desbalanceada.

O banco de dados contém diversas variáveis que podem explicar a possível inadimplência, considerando as características do tomador e da operação como, por exemplo, (i) renda anual, (ii) tipo de imóvel (próprio, alugado, financiado,...), (iii) motivo do empréstimo (pessoal, educação, saúde, investimento em negócios, reforma de imóvel, ...), (iv) taxa de juros da operação, (v) *rating* inicial da operação, etc. Após a realização de uma limpeza da base, eliminando-se dados faltantes e agregando-se variáveis nominais com poucas observações, obteve-se uma base final com 28.544 operações de empréstimo.

O segundo banco de dados baseia-se em levantamento de Moro, Cortez e Rita (2014) composto 45.211 observações que refletem campanhas de *marketing* direto em uma instituição financeira de Portugal. A variável a ser classificada relaciona-se com o cliente aderir ou não aderir a um programa de marketing e realizar depósitos no banco. Nesse problema específico, deseja-se avaliar se um dado cliente é mais ou menos susceptível a esforços de contato voltados para oferta de um produto financeiro.

Na base de dados, a taxa de conversão, ou seja, a taxa de clientes que fazem os depósitos bancários a partir dos esforços de *marketing*, é de cerca de 12% indicando um desbalanceamento maior que a primeira base de dados. É importante ressaltar que o desbalanceamento da frequência de observações entre as diferentes classes constitui um desafio a modelos quantitativos de classificação baseados em aprendizagem de máquina.

As variáveis explicativas da segunda base de dados envolvem, por exemplo, (i) idade do cliente, (ii) tipo de emprego (administrativo, operário, empreendedor, estudante, autônomo, ...), (iii) estado civil, (iv) nível educacional, (v) forma de contato (telefone ou celular), (vi) dia da semana do último contato, (vii) duração do contato em segundos, (viii) resultado de campanhas anteriores para esse cliente, (ix) taxa de desemprego, (x) taxa de inflação, etc...

Em todos os bancos de dados, a amostra de treinamento foi escolhida obtendo-se observações aleatórias correspondentes a 70% da amostra total. Uma vez treinados os modelos de aprendizagem de máquina, os parâmetros obtidos são utilizados para fazer a previsão na amostra de teste, que corresponde ao restante 30% da amostra total. Foi realizada validação cruzada em 10-fold, repetindo 3 vezes.

Os algoritmos foram rodados em R, utilizando bibliotecas com funções *glm*, *treebag*, *gbm* e *rf* que exploram algoritmos de regressão logística, *bagging*, *boosting* e *random forests*, respectivamente. Para o *bagging*, a técnica de classificação base é a árvore de decisão e no *boosting* foi usado o *stochastic gradient boosting*, baseado em árvores.

Para avaliação de adequação dos modelos na amostra de teste foram consideradas diversas métricas baseadas na matriz de confusão. Ressalta-se que a matriz de confusão indica o acerto ou não de uma predição realizada pelos algoritmos, a partir das variáveis explicativas, em relação ao verdadeiro valor da variável de interesse.

A partir da notação de *TP* para positivos verdadeiros, *FN* para falsos negativos, *FP* para falsos positivos e *TN* para verdadeiros negativos, foram calculados (i) a sensibilidade ou taxa de verdadeiros positivos, também chamada de *recall* ($S = TP / (TP + FN)$), (ii) a especificidade

ou taxa de verdadeiros negativos ($E = TN/(TN + FP)$), (iii) acurácia ($A = (TP + TN)/(TP + TN + FP + FN)$), (iv) valor preditivo negativo ($NPV = TN/(TN + FN)$), (v) escore F1 ($F1 = 2S \cdot PPV/(S + PPV)$), (vi) prevalência: ($PV = (TP + FN)/(TP + TN + FP + FN)$), (vii) taxa de detecção ($DR = TP/(TP + TN + FP + FN)$), (viii) taxa de prevalência ($DP = (TP + FP)/(TP + TN + FP + FN)$), (ix) acurácia balanceada ($BA = (S + E)/2$). Os resultados são apresentados nas Tabelas 1 e 2.

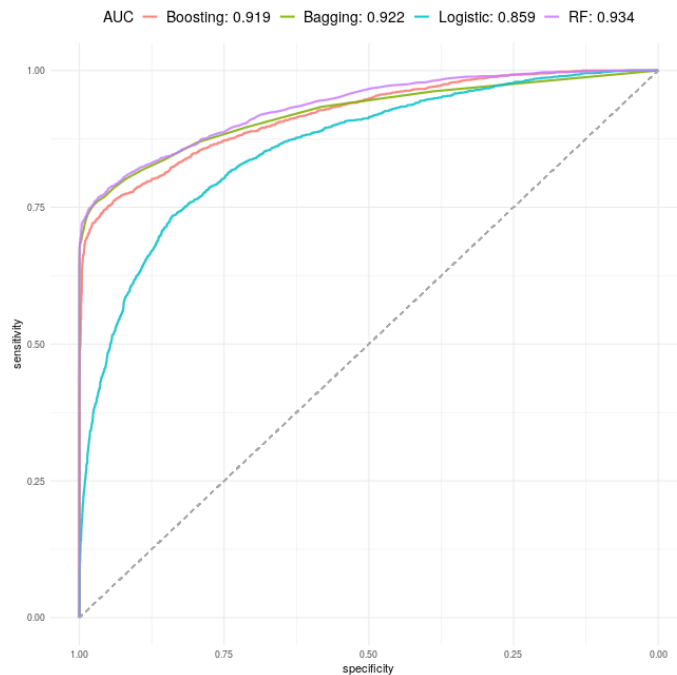
Cada métrica focaliza o desempenho do modelo em termos de maior número de acertos na previsão, considerando diferentes dimensões que podem ser avaliadas a partir da matriz de confusão. Foram também geradas as curvas ROC (*Receiver Operating Characteristic*), ilustradas nas Figuras 1 e 2, que permitem uma avaliação visual da adequação de cada modelo.

Tabela 1: Métricas de desempenho de classificação dos diversos modelos para o banco de dados de empréstimos para pessoa física

Desempenho	Regressão Logística	Bagging	Boosting	Random Forest
S	0,947	0,986	0,983	0,992
E	0,495	0,736	0,704	0,729
A	0,873	0,932	0,924	0,931
NPV	0,719	0,934	0,919	0,961
F1	0,909	0,958	0,953	0,960
PV	0,786	0,786	0,786	0,786
DR	0,744	0,774	0,772	0,779
DP	0,852	0,831	0,836	0,837
BA	0,721	0,861	0,843	0,860
AUROC	0,859	0,922	0,919	0,934

Os resultados indicam que para a primeira base de dados, *random forest* e *bagging* são superiores aos outros modelos em diversas métricas de desempenho. Em particular, o *random forest*, método *ensemble* baseado na combinação de árvores de decisão, mostra um desempenho superior em diversas métricas, inclusive considerando a curva ROC e a métrica AUROC.

Figura 1: Curva ROC dos modelos aplicados à base de inadimplência



Evidencia-se que a melhoria de desempenho em relação à técnica tradicional de classificação baseada na regressão logística é substancial, principalmente quando se leva em consideração que o ganho de acurácia envolve praticamente apenas esforço computacional, sem a necessidade de uma atuação de um analista de crédito humano.

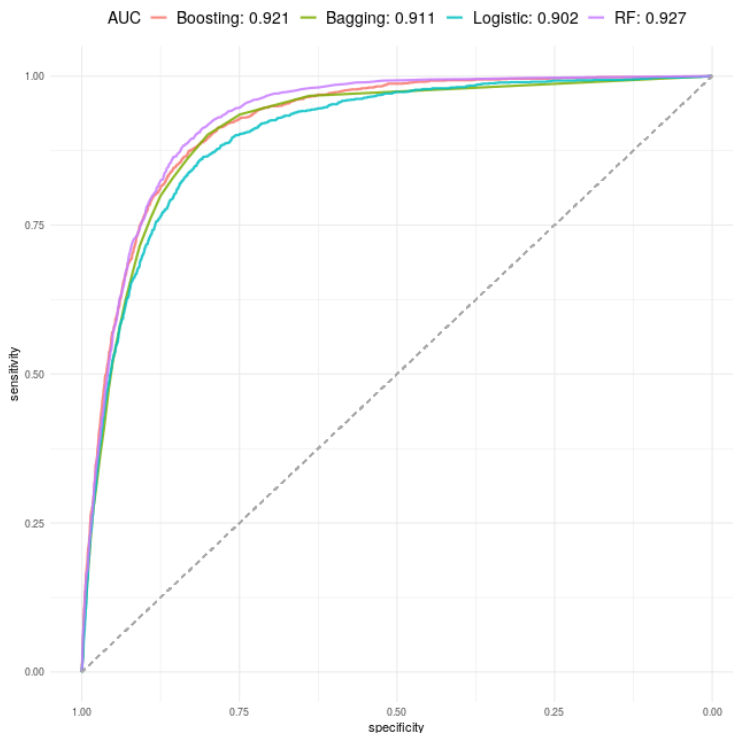
Já para a segunda base de dados, referente à sensibilização do cliente a esforços de *marketing* da instituição financeira, os resultados são menos decisivos para uma técnica específica. Diferentes métricas de performance favorecem diferentes métodos. Dependendo dos objetivos de assertividade da instituição financeira, os métodos *ensemble* de *bagging* e *boosting* e até mesmo a técnica estatística tradicional de regressão logística podem servir de modelo para tomada de decisão.

Tabela 2: Métricas de desempenho de classificação dos diversos modelos para o banco de dados de adesão a produto de depósito bancário.

Desempenho	Regressão Logística	Bagging	Boosting	Random Forest
S	0,974	0,957	0,971	0,964
E	0,349	0,473	0,417	0,462
A	0,920	0,933	0,927	0,932
NPV	0,639	0,592	0,651	0,623
F1	0,946	0,945	0,949	0,948
PV	0,885	0,885	0,885	0,885
DR	0,862	0,847	0,859	0,852
DP	0,937	0,908	0,926	0,914
BA	0,661	0,715	0,694	0,713
AUROC	0,902	0,911	0,921	0,927

Enquanto na base anterior o *random forest* propiciava melhores classificações em uma série de métricas, nessa segunda base, a técnica de combinação de árvores de decisão apresenta superioridade apenas em termos do indicador de AUROC.

Figura 2: Curva ROC dos modelos aplicados à base de adesão a esforços de *marketing*



5. Considerações finais

Sob uma perspectiva teórica, as inovações em finanças buscam explorar oportunidades associadas a imperfeições de mercado, custos de transação ou lacunas regulatórias. Sob uma perspectiva prática, as inovações financeiras podem trazer benefícios para uma gama enorme de *stakeholders*. Não somente agentes que buscam obter aumento de riqueza podem se beneficiar das inovações como também a sociedade pode se desenvolver.

Por exemplo, conforme discutido nesse estudo, as inovações em produto e em serviços financeiros podem propiciar (i) o financiamento de atividades de alto grau de risco e de iniciativas de impacto social, que não necessariamente envolvem uma busca direta por lucro, bem como (ii) o acesso de indivíduos ao mercado bancário e a instrumentos financeiros. Neste contexto, o desenvolvimento de um mercado de carbono e o surgimento de *fintechs* constituem exemplos de inovação que podem contribuir para a sociedade em geral.

No entanto, apesar de seus potenciais benefícios, as inovações também podem trazer riscos ao mercado. Exemplificando, (i) a inovação em produto com a criação de instrumentos financeiros sofisticados e (ii) o uso de algoritmos automatizados de *high frequency* estão associados à crise de crédito de 2007-2008 e ao flash crash do mercado acionário americano em 2010, respectivamente.

Neste contexto, o surgimento de uma nova onda de inovações financeiras propiciada pelo avanço tecnológico, mais especificamente, em relação ao levantamento, armazenamento e processamento de dados e ao desenvolvimento de algoritmos de inteligência artificial, deve ser visto com otimismo e também com cautela.

Do ponto de vista positivo, as inovações financeiras podem amplificar benefícios à sociedade, reduzindo custos de transação e propiciando maior acesso da população aos instrumentos e mercados financeiros. Porém, do ponto de vista negativo, desafios surgem com relação a questões de privacidade, de proteção de dados, de uso ético da inteligência artificial, etc.

Nesse estudo, buscou-se evidenciar contribuições de inteligência artificial para a indústria bancária, com foco portanto na perspectiva otimista de que técnicas de aprendizagem de máquina podem trazer benefícios ao segmento bancário. Foram apresentadas aplicações de inteligência artificial em problemas de classificação notadamente enfrentados por bancos: (i) avaliação de risco de crédito, buscando-se gerar modelos de classificação de bons e maus pagadores em operações de crédito e (ii) análise de esforços de *marketing*, buscando-se obter modelos que melhor identifiquem clientes propensos ou susceptíveis a adquirir produtos financeiros.

A partir da aplicação de métodos *ensemble* de aprendizagem de máquina, mais especificamente, *bagging*, *boosting* e *random forest*, foram comparados os desempenhos de métricas de classificação com uma técnica estatística usualmente utilizada como *benchmark*, a regressão logística.

É importante destacar que os resultados apresentados não permitem explorar a importância relativa das variáveis explicativas para previsão da variável dependente, evidenciando a crítica comumente realizada a técnicas de aprendizagem de máquina sobre representarem *black boxes*. Embora existam ferramentas para identificar a relevância de cada variável para a previsão, esses

mecanismos não foram explorados nesse estudo. Além disso, análises de balanceamento e outras técnicas de validação não foram analisadas.

Como sugestão para estudos futuros, podem ser indicados (i) o estudo do desbalanceamento dos bancos de dados, (ii) a análise de seleção de características ou variáveis que poderiam trazer maiores ganhos de classificação, (iii) a avaliação de desempenho ajustada por erros de classificação, ou seja, considerando maior peso para clientes que são considerados bons pagadores mas que deveriam ser classificados com mau pagadores e para clientes considerados susceptíveis aos esforços de marketing, mas que deveriam ser classificados como não susceptíveis. Tendo em vista que custos de inadimplência e custos dos esforços de *marketing* podem ser substanciais, atribuir pesos diferentes a classificações erradas pode auxiliar no processo de tomada de decisão.

Finalmente, pode-se indicar que as inovações tecnológicas, notadamente as advindas pelo uso de inteligência artificial, devem ser vistas sob oportunidades e ameaças. O uso de IA pode melhorar a experiência do cliente, reduzir custos de transação e dar acesso a produtos e serviços financeiros a uma maior parcela da população. Em contrapartida, confiar demasiadamente em resultados de modelos de inteligência artificial, sem um domínio claro de premissas ou de atendimento a regras éticas e morais, pode implicar em violação de privacidade e aumento de discriminação.

Dessa forma, instituições financeiras, agentes reguladores, profissionais de finanças e demais *stakeholders* devem compreender os avanços em finanças e avaliar de forma técnica as inovações que a inteligência artificial tem trazido, buscando aumentar os benefícios à sociedade e limitar os impactos negativos ao mercado.

Referências bibliográficas

Accenture. From bottom line to front line. **The CFO Reimagined: from Driving Value to Building the Digital Enterprise**. Relatório Técnico. Accenture, 2018a. Disponível em: https://www.accenture.com/_acnmedia/PDF-85/Accenture-CFO-Research-Global.pdf#zoom=40. Acesso em: 12/03/2020.

Aggarwal, Divya; Chandrasekaran, Shabana; Annamalai, Balamurugan. A Complete Empirical Ensemble Mode Decomposition and Support Vector Machine-Based Approach to Predict Bitcoin Prices. **Journal of Behavioral and Experimental Finance**, v. 27, 100335, 2020.

Aliber, Robert. Financial Innovation and the Boundaries of Banking. **Managerial and Decision Economics**, v. 8, n. 1, p. 67-73, 1987.

Allen, Franklin. Trends in Financial Innovation and their Welfare Impact: an Overview. **European Financial Management**, v. 18, n. 4, p. 493-514, 2012.

Awalegaonkar, Ketan; Berkey, Robert; Douglass, Greg; Reilly, Athena. AI: Built to Scale: from Experimental to Exponential. Relatório Técnico. **Accenture**, 2019. Disponível em: https://www.accenture.com/_acnmedia/Thought-Leadership-Assets/PDF-2/Accenture-Built-to-Scale-PDF-Report.pdf#zoom=50. Acesso em 12/03/2020.

Avgouleas, Emiliós. Regulating Financial Innovation: A Multifaceted Challenge to Financial Stability, Consumer Protection and Growth. In Moloney, Niamh; Ferran, Eilis; Payne, Jennifer (eds), **Oxford Handbook on Financial Regulation**. Oxford: Oxford University Press, 2015.

Babel, Bernhard; Buehler, Kevin; Pivonka, Adam; Richardson, Bryan; Waldron, Derek. Derisking Machine Learning and Artificial Intelligence. **McKinsey on Risk**, n. 7, fevereiro, 2019. Disponível em:

<https://www.mckinsey.com/~media/McKinsey/Business%20Functions/Risk/Our%20Insights/Derisking%20machine%20learning%20and%20artificial%20intelligence/Derisking-machine-learning-and-artificial-intelligence-vF.ashx>. Acesso em 03/04/2020.

Beutel, Johannes; List, Sophia; von Schweinitz, Gregor. Does machine learning help us predict banking crises? *Journal of Financial Stability*, v. 45, 100692, 2019.

Bavoso, Vincenzo. Financial Innovation, Derivatives and the UK and US Interest Rate Swap Scandals: Drawing New Boundaries for the Regulation of Financial Innovation. **Global Policy**, v. 7, n. 2, p. 227-236, 2016.

Benston, George; Smith, Jr., Clifford. A Transaction Cost Approach to the Theory of Financial Intermediation. **Journal of Finance**, v. 31, p. 215-231, 1976.

BIS. Big tech in finance: opportunities and risks. BIS Annual Economic Report. **Bank of International Settlements**. 23 de junho de 2019. Disponível em: <https://www.bis.org/publ/arpdf/ar2019e3.pdf>. Acesso em: 22/04/2020.

Black, Fischer; Scholes, Myron. From Theory to New Financial Product. **Journal of Finance**, v. 41, p. 645-655, 1986.

Blake, David. Financial Intermediation and Financial Innovation in a Characteristics Framework. **Scottish Journal of Political Economy**, v. 43, p. 16-31, 1996.

Boser, Bernhard; Guyon, Isabelle; Vapnik, Vladimir. A training algorithm for optimal margin classifiers. **Proceedings of the fifth annual workshop on Computational learning theory**, p. 144-152, 1992.

Brainard; Lael. What Are We Learning about Artificial Intelligence in Financial Services? **Fintech and the New Financial Landscape**, Philadelphia, Pennsylvania. novembro, 2018. Disponível em: <https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm>. Acesso em: 15/02/2020.

Breiman, Leo. Bagging predictors. Department of Statistics, **University of California Berkeley**, Technical Report 421, 1994.

Breiman, Leo. Bagging predictors. **Machine Learning**, v. 24, n. 2, p. 123-140, 1996.

Breiman, Leo. Random forests. **Machine Learning**, v. 45, n. 1, p. 5-32, 2001.

Callon, Michel.; Lacoste, Annalivia, Defending Responsible Innovation. **Debating Innovation**, v. 1, n. 1, p. 19–27, 2011.

Caputo, Victor. Itaú usará inteligência artificial para ter agências mais seguras. **Exame**, Tecnologia. 12 de julho de 2017a. Disponível em:
<https://exame.abril.com.br/tecnologia/itau-usara-inteligencia-artificial-para-ter-agencias-mais-seguras/>. Acesso em 13/4/2020.

Caputo, Victor. Bradesco usa inteligência artificial no atendimento a clientes. **Exame**, Tecnologia. 19 de julho de 2017b. Disponível em:
<https://exame.abril.com.br/tecnologia/bradesco-usa-inteligencia-artificial-no-atendimento-a-clientes/> Acesso em: 13/04/2020.

Columbus, Louis. 10 Ways AI Is Going To Improve Fintech In 2020. **Forbes**, Enterprise & Cloud. 29 de dezembro de 2019. Disponível em:
<https://www.forbes.com/sites/louiscolumbus/2020/12/29/10-ways-ai-is-going-to-improve-fintech-in-2020/#c98819d4ffd2>. Acesso em: 12/03/2020.

du Jardin, Philippe. Forecasting Corporate Failure Using Ensemble of Self-Organizing Neural Networks. **European Journal of Operational Research**, in proof. Disponível online em 19/06/2020

Faustino, Rafael. Como a Inteligência Artificial Está Mudando o Sistema Financeiro. **Época**, Negócios. 17 de agosto de 2018. Disponível em:
<https://epocanegocios.globo.com/Dinheiro/noticia/2018/08/como-inteligencia-artificial-esta-mudando-o-sistema-financeiro.html>. Acesso em: 21/04/2020.

- Fama, Eugene. Agency Problems and the Theory of the Firm. **Journal of Political Economy**, v. 88, n. 2, p. 288-307, 1980.
- Freund, Yoav. Boosting a weak learning algorithm by majority. **Information and Computation**, v. 121, n. 2, p. 256-285, 1995.
- Freund, Yoav; Schapire, Robert. Experiments with a new boosting algorithm. **Machine Learning: Proceedings of the Thirteenth International conference**, p. 148-156, 1996.
- Freund, Yoav; Schapire, Robert. A short introduction to boosting. **Journal of Japanese Society for Artificial Intelligence**, v. 14, n. 5, p. 771-780, 1999.
- Frieden, Jeffrey. **Global Capitalism, its Fall and Rise in the Twentieth Century**. New York, NY: Norton, 2006.
- Gezer, Ali; Wanrner, Gary; Wilson, Clifford; Shrestha, Prakash. A flow-base approach for Trickbot banking trojan detection. **Computers & Security**, v. 84, p. 179-192, 2019.
- Ghahramani, Zoubin. Probabilistic Machine Learning and Artificial Intelligence. **Nature**, v. 521, 452-459, 2015.
- Golbayani, Parisa; Florescu, Ionut; Chatterjee, Rupak. A comparative study of forecasting credit ratings using neural networks, support vector machines, and decision trees. **The North American Journal of Economics and Finance**, v. 54, 101251, 2020.
- Israel, González-Carrasco; Jose, Jiménez-Márquez, Jose, López-Quadrado; Belén, Ruiz-Mescia. Automatic detection of relationships between banking operations using machine learning. **Information Sciences**, v. 485, p. 319-346, 2019.
- Grinblatt, Mark; Longstaff, F. A. Financial Innovation and the Role of Derivative Securities: An Empirical Analysis of the Treasury STRIPS Program. **Journal of Finance**, v. 55, n. 3, p. 1415-1436, 2000.

- Henderson, Brian; Pearson, Neil. The Dark Side of Financial Innovation: a Case Study of the Pricing of a Retail Financial Product. **Journal of Financial Economics**, v. 100, p. 227–247, 2011.
- Heric, Michael. Rethinking How Finance Uses Digital Tools. Relatório Técnico. **Bain & Co**, 2018. Disponível em:
https://www.bain.com/contentassets/9cd677fc720d4e28b2ea4a6076190afa/bain_brief_rethinking_how_finance_uses_digital_tools.pdf. Acesso em: 10/03/2020.
- Ho, Tin. Random decision forests. **Proceedings of 3rd International Conference on Document Analysis and Recognition**, v. 1, n. 1, p. 278-282, 1995.
- Immergluck, Dan. Core of the Crisis: Deregulation, the Global Savings Glut, and Financial Innovation in the Subprime Debacle. A Symposium on the Subprime Crisis. **City & Community**, v. 8, n. 3, p. 341-356, 2009.
- Jordan, Steven; Vivian, Andrew; Wohar, Mark. Forecasting Market Returns: Bagging or Combining? **International Journal of Forecasting**, v. 33, p. 102-120, 2017.
- Jubraj, Roy; Graham, Tom; Ryan, Eve. Redefine banking with Artificial Intelligence. Relatório Técnico. **Accenture**, 2018b. Disponível em:
https://www.accenture.com/_acnmedia/pdf-68/accenture-redefine-banking.pdf. Acesso em: 20/03/2020.
- KPMG. Living in an AI World. Financial Services Report. Relatório Técnico. **KPMG**, 2020. Disponível em: <https://advisory.kpmg.us/content/dam/advisory/en/pdfs/2020/financial-services-living-in-ai-world.pdf>. Acesso em: 19/04/2020.
- Krauss, Christopher; Do, Xuan; Huck, Nicolas. Deep Neural Networks, Gradient-Boosted Trees, Random Forests: Statistical Arbitrage on the S&P 500. **European Journal of Operational Research**, v. 259, n. 2, p. 689-702, 2017.

- Kuncheva, Ludmila; Whitaker, Christopher. Measures of Diversity in Classifier Ensembles, **Machine Learning**, v. 51, p. 181-207, 2003
- Ladyzynski, Piotr; Zbikowski, Kamil; Gawrysiak, Piotr. Direct marketing campaigns in retail banking with the use of deep learning and random forests. *Expert Systems With Applications*, v. 134, p. 28-35, 15/11/2019.
- Liang, Deron; Tsai, Chih-Fong; Lu, Hung-Yuan; Chang, Li-Shin. Combining Corporate Governance Indicators with Stacking Ensembles for Financial Distress Prediction. **Journal of Business Research**, v. 120, p. 137-146, 2020.
- Mari, Angelica. A Aposta do Bradesco em Inteligência Artificial. **Forbes**, Negócio, 31 de março de 2019. Disponível em: <https://forbes.com.br/negocios/2019/03/a-aposta-do-bradesco-em-inteligencia-artificial/>. Acesso em: 14/04/2020.
- Marzochi, Roger. Em casa, com inteligência. **Noomis**, Ciab Febraban. 30 de março de 2018. Disponível em: https://noomis.febraban.org.br/temas/inteligencia-artificial/em-casa-com-inteligencia?utm_source=related_content. Acesso em: 21/04/2020.
- McCauley, Denis. The Global AI Agenda: Promise, Reality, and a Future of Data Sharing. Technical Report. **MIT Technology Review Insights**, 2020. Disponível em: <https://mittrinsights.s3.amazonaws.com/AIagenda2020/GlobalAIagenda.pdf>. Acesso em: 19/04/2020.
- McWaters, Jesse; Galaski, Rob. The New Physics of Financial Services. Understanding How Artificial Intelligence is Transforming the Financial Ecosystem. **World Economic Forum e Deloitte**. Relatório Técnico, agosto de 2018. Disponível em: http://www3.weforum.org/docs/WEF_New_Physics_of_Financial_Services.pdf. Acesso em 19/04/2020.
- Miller, Merton. Financial Innovation: Achievements and Prospects. **Journal of Applied Corporate Finance**, v. 4, n. 4, p. 4-11, 1992.

- Millo, Yuval; MacKenzie, Donald. The Usefulness of Inaccurate Models: Towards an Understanding of the Emergence of Financial Risk Management. **Accounting, Organizations and Society**, v. 34, n. 5, p. 638–653, 2009.
- Modigliani, Franco; Miller, Merton. The Cost of Capital, Corporation Finance and the Theory of Investment. **American Economic Review**, v. 48, n. 3, p. 261-297, 1958.
- Modigliani, Franco; Miller, Merton. Corporate Income Taxes and the Cost of Capital: a Correction. **American Economic Review**, v. 53, n. 3, p. 433-443, 1963.
- Moro, Sérgio; Cortez, Paulo; Rita, Paulo. A data-driven approach to predict the success of bank telemarketing. **Decision Support Systems**, v. 62, p. 22-31, 2014.
- Muniesa, Fabian; Lenglet, Marc. Responsible Innovation in Finance: Directions and Implications. In Owen, Richard; Bessant, John; Heintz, Maggy (Eds). **Responsible Innovation**. John Wiley & Sons. Abril, 2013.
- Optiz, David; Maclin, Richard. Popular Ensemble Methods: an Empirical Study. **Journal of Artificial Intelligence Research**, v. 11, p. 169-198.
- Pantano, Eleonora; Pizzi, Gabriele. Forecasting artificial intelligence on online customer assistance: Evidence from chatbot patents analysis. **Journal of Retailing and Consumer Services**, v. 55, 102906, 2020.
- Plaschke, Frank; Seth, Ishaan; Whiteman, Rob. Bots, Algorithms, and the Future of the Finance Function. **McKinsey on Finance**, n. 65, January 2018. Disponível em: <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Strategy%20and%20Corporate%20Finance/Our%20Insights/Bots%20algorithms%20and%20the%20future%20of%20finance%20function%20new/Bots-algorithms-and-the-future-of-the-finance-function.ashx>. Acesso em: 15/04/2020.
- Reuters. BC Regulamenta Fintechs e Define Limite de R\$ 15 mil para Empréstimos. Estadão Conteúdo. 27 de abril de 2018. Disponível em: <https://veja.abril.com.br/economia/bc->

- regulamenta-fintechs-e-define-limite-de-r-15-mil-para-emprestimos/. Acesso em: 24/04/2020.
- Rokach, Lior. Ensemble-based Classifiers. **Artificial Intelligence Review**, v. 33, n. 1–2, p. 1-39, 2010.
- Romani, Bruno. PicPay faz Aposta em Inteligência Artificial. **Estadão**, Link. 11 de março de 2020. Disponível em: <https://link.estadao.com.br/noticias/inovacao,picpay-faz-aposta-em-inteligencia-artificial,70003227859>. Acesso em: 20/04/2020.
- Ross, Stephen. Institutional Markets, Financial Marketing, and Financial Innovation. **Journal of Finance**, v. 44, n. 3, p. 541-556, 1989.
- Rtayli, Naoufal; Enneya, Nourddine. Selection features and support vector machine for credit card risk identifications. **Procedia Manufacturing**, v. 46, p. 941-948, 2020.
- Shevlin, Ron. Amazon’s Impending Invasion of Banking. **Forbes**, Fintech. 8 de julho de 2019. Disponível em: <https://www.forbes.com/sites/ronshevlin/2019/07/08/amazon-invasion/#1050b3d97921>. Acesso em: 22/04/2020.
- Sollich, Peter; Krogh, Anders. Learning with ensembles: How Overfitting can be Useful. **Advances in Neural Information Processing Systems**, v. 8, p. 190-196, 1996.
- Swaroop, Himanshu Artificial Intelligence in Finance: Five Opportunities to Take the Leap. Blog. **Accenture**, 4 de setembro de 2019. Disponível em: <https://financeandriskblog.accenture.com/finance-accounting/artificial-intelligence-in-finance-five-opportunities-to-take-the-leap>. Acesso em: 15/04/2020.
- Taleb, Nassim. **The Black Swan: The Impact of the Highly Improbable**. Random House, New York, 2007.
- Tse, Lao. Credit Risk Dataset. Banco de dados. **Kaggle**. Disponível em: <https://www.kaggle.com/laotse/credit-risk-dataset/metadata>. Acesso em: 10/08/2020.

- Tufano, Peter. Financial Innovation. Constantinides, George; Harris, Milton; Stulz, René (eds). **Handbook of the Economics of Finance**, v.1, parte A, p. 307-335, 2003.
- Vieira, José; Barboza, Flavio; Sobreiro, Vinicius; Kimura, Herbert. Machine learning models for credit analysis improvements: Predicting low-income families' default. **Applied Soft Computing**, v. 83, 105640, outubro, 2019.
- Wang, Jianzhou; Niu, Tong;, Du,, Pei; Yang,, Wendong. Ensemble Probabilistic Prediction Approach for Modeling Uncertainty in Crude Oil Price. **Applied Soft Computing**, v. 95, 106509, 2020.
- Wieneke, Axel. Better Financial Innovation via Innovative Finance of Supervisors. **Economic Papers: A Journal of Applied Economics and Policy**, v. 35, n. 1, p. 16-23, 2015.
- Yin, Anwen. Equity Premium Prediction and Optimal Portfolio Decision with Bagging. **The North American Journal of Economics and Finance**, v. 54, 101274, 2020.
- Yu, Howard; Rohner, Urs. A Inteligência Artificial Vai Chacoalhar o Setor Bancário. **Exame**, Tecnologia. 14 de dezembro de 2018. Disponível em:
<https://exame.abril.com.br/tecnologia/a-inteligencia-artificial-vai-chacoalhar-o-setor-bancario/>. Acesso em: 21/04/2020.
- Zięba, Maciej; Tomczak, Sebastian; Tomczak, Jakub. Ensemble Boosted Trees with Synthetic Features Generation in Application to Bankruptcy Prediction. **Expert Systems With Applications**, v. 58, p. 93-101, 2016.