O crédito é necessário para financiar diversos iniciativas e negócios. Recentemente, credores e empresas de tecnologia financeira (“fintech”) estão procurando usar fontes alternativas de dados e formas inovativas de análise de risco de crédito. Essas inovações podem ampliar o acesso ao crédito, especialmente para pessoas com históricos de crédito escassos. No mercado financeiro do consumidor, dados alternativos referem-se a informações usadas para avaliar a qualidade de crédito que geralmente não fazem parte de um relatório de crédito.

Alguns exemplos incluem:

* Pagamentos de aluguel.
* Pagamentos por celular.
* Pagamentos de TV a cabo.
* Sites de reputação on-line.
* Métricas de análise de sentimento de mídias sociais.

O uso de dados alternativos tem o potencial de ajudar a expandir o acesso responsável ao crédito. Por exemplo, alguém sem histórico anterior de pagamento de empréstimo ainda terá variabilidade em tais informações alternativas. Outra vantagem é disponibilizar escores alternativos de risco de crédito.

As políticas contratuais de alguns credores geralmente exigem o empréstimo seja efetuado para empresas com escores de crédito acima de determinado limiar, mas sem restringir que sejam escores de *bureaus* de crédito tradicionais. Alguns desses credores podem estar dispostos a conceder o empréstimo, se encontrarem critérios objetivos para determinar quais pessoas têm menos probabilidade de inadimplência no empréstimo, observando para isto outras fontes de dados.

Os dados tradicionalmente usados pelos credores e *bureaus* de crédito, podem não refletir todas as atividades que a empresa se envolve. Dados alternativos podem fornecer informações mais atualizadas e em tempo real.

Informações imprecisas ou incompletas. Embora os dados tradicionais também possam ser imprecisos, certos tipos de dados alternativos podem ser mais propensos a erros se os padrões que regem esses dados forem diferentes ou mais fracos do que aqueles que regem os dados tradicionais. Os consumidores podem não conseguir acessar ou visualizar alguns tipos de dados alternativos. Isso pode impedir que os consumidores encontrem e corrijam quaisquer imprecisões.

Dificuldades em alcançar uma boa posição. Os fatores de crédito tradicionais são fortemente influenciados pela própria conduta financeira de uma pessoa. Alguns dados alternativos podem não estar relacionados à própria conduta financeira de uma pessoa e o uso desses dados pode tornar mais difícil para as pessoas melhorarem sua posição de crédito. Fatores alternativos de crédito também podem ser mais difíceis de explicar para as pessoas que procuram crédito.

Efeitos colaterais não intencionais. O uso de dados alternativos pode penalizar ou recompensar certos grupos ou comportamentos de maneiras difíceis de prever. Por exemplo, militares podem se mudar com frequência e isso pode dar uma falsa impressão de instabilidade que pode afetar seu acesso ao crédito.

O potencial de discriminação. O uso de dados alternativos pode apresentar um risco maior de discriminação ilegal se novas variáveis ou fatores estiverem mais relacionados a um fator que não pode ser usado de acordo com a lei (como raça, etnia ou gênero).

Como as pessoas podem saber quando dados alternativos estão sendo usados para tomar decisões sobre eles no mercado financeiro do consumidor?

Boa pergunta. Os credores podem ne

Com a introdução das diretrizes de conformidade de Basel II e Basel III, e a necessidade de mais acurácia nas mensurações de risco de crédito, os modelos de análise de sobrevivência ganharam maior importância ao longo dos anos (Dirick, 2012). Historicamente, a análise de sobrevivência tem sido utilizada principalmente em medicina e engenharia, onde o tempo de duração até ocorrência de um evento é analisada.

Narain (1992) foi a primeira pessoa que sugeriu a utilização de análise de sobrevivência no contexto de risco de crédito. A vantagem da utilização

Time to Default in Credit Scoring Using Survival Analysis: A benchmark study. Journal of The Operational Research Society, v. 68, p. 652-665.

SHAP e LIME: Uma Avaliação do Poder Discriminativo no Risco de Crédito

Esses modelos são geralmente chamados de “caixas-pretas”, o que significa que você conhece as entradas e saídas, mas há poucas maneiras de entender o que está acontecendo sob o capô. Como resposta a isso, vimos vários modelos diferentes de IA explicada florescerem nos últimos anos, com o objetivo de permitir que o usuário veja por que a caixa preta deu uma determinada saída. Nesse contexto, avaliamos dois modelos eXplainable AI (XAI) muito populares em sua capacidade de discriminar observações em grupos, por meio da aplicação de modelagem não supervisionada e preditiva aos pesos que esses modelos XAI atribuem a recursos localmente. A avaliação é realizada em dados reais de Pequenas e Médias Empresas, obtidos de repositórios oficiais italianos, e pode servir de base para o emprego de tais modelos XAI para extração de características pós-processamento.

1. Introdução

A estimativa da probabilidade de inadimplência (PD) é uma questão que os bancos e outras instituições financeiras têm enfrentado desde o surgimento do crédito. Sistemas e metodologias evoluíram junto com o conhecimento e a tecnologia, mas não foi até recentemente que os incríveis avanços dados em TI deram uma verdadeira sacudida na forma como ela era executada pela indústria. A princípio, as instituições incumbentes resistiram à aplicação de novos paradigmas, o que favoreceu o surgimento de um número crescente de startups de fintech cujo objetivo é fornecer uma estimativa da solvabilidade de pessoas e empresas e fazer com que essa estimativa seja a mais alta fidelidade possível.

Para poder fornecer essa estimativa, é claro que essas empresas alavancam novas e diversas fontes de dados, aproveitam as inovações na estrutura regulatória relativa a dados financeiros (por exemplo, PSD2 europeu (Comissão Europeia, (2015)) e exploram o poder preditivo muito maior que alguns dos algoritmos recém-implementados oferecem respeito aos métodos tradicionais. No entanto, o aumento no poder de previsão dos novos algoritmos prejudica a explicabilidade, uma vez que os modelos agora são tão complexos que é quase impossível estabelecer vínculos claros entre os funcionamento do modelo e do output dado, o que seguramente representa um problema e dificulta a sua difusão, para além de levantar uma série de questões éticas e regulamentares, que começam a ser abordadas (ver, por exemplo, Comissão Europeia (2020)).

Para resolver esse trade-off, o conceito de eXplainable AI (XAI) surgiu introduzindo um conjunto de técnicas de aprendizado de máquina (ML) que produzem modelos que oferecem um trade-off aceitável entre explicabilidade e utilidade preditiva e permite que os humanos entendam, confiem e gerenciar as gerações emergentes de modelos de IA. Entre as técnicas emergentes, dois frameworks têm sido amplamente reconhecidos como o estado da arte em eXplainable AI e são eles:

• o framework Lime, introduzido por (Ribeiro et al., 2016)

• Valores SHAP, introduzidos por (Lundberg e Lee, 2017).

Em finanças, a interpretabilidade é especialmente importante porque a confiança do modelo nas características corretas deve ser garantida; ainda, não existem muitos estudos com foco na aplicação do XAI neste contexto específico (Bussmann, 2020). propõem um modelo XAI baseado em valores de Shapley aplicado no contexto de decisões de empréstimo para PME que buscam financiamento por meio de plataformas P2P, enquanto a pesquisa de (Ariza-Garzón et al., 2020) visa avaliar a capacidade preditiva de vários modelos de ML no contexto de pontuação de crédito das plataformas de empréstimo P2P, em seguida aplicando o método Shapley para fornecer explicabilidade à previsão. O precedente mais interessante talvez seja a pesquisa de (HadjiMisheva et al., 2021), onde os autores exploram a utilidade das estruturas SHAP e Lime no contexto do gerenciamento de risco de crédito, delineando os obstáculos práticos na aplicação dessas técnicas a vários tipos diferentes de algoritmos de ML, bem como propor soluções para os desafios enfrentados.

Nosso estudo visa comparar as estruturas SHAP e LIME avaliando sua capacidade de definir grupos distintos de observações, empregando os pesos atribuídos às feições por meio de seu algoritmo de interpretabilidade local como espaço de entrada para abordagens não supervisionadas e supervisionadas. Fazemos isso construindo nossa abordagem em um dos algoritmos de aprendizado supervisionado de melhor desempenho, porém complexo, XGBoost (Chen e Guestrin, 2016), empregado para prever a probabilidade de inadimplência de pequenas e médias empresas italianas.