

Precificação de Dívidas Orientada a Dados: Revisão Sistemática da Literatura

Bryan L. M. de Oliveira¹, Telma W. de L. Soares¹, Wellington S. Martins¹

¹ Instituto de Informática – Universidade Federal de Goiás (UFG)

{bryanlincoln,telma,wellington}@inf.ufg.br

Abstract. *The Debt Collection (RC) market presents vast potential for enhancing its processes through the implementation of Machine Learning (AM) techniques, due to the profusion of data accumulated by debt collection institutions. This Systematic Literature Review aims to identify techniques applied to the specific predicament of Debt Pricing (PD), especially Reinforcement Learning, and to highlight research opportunities that remain unexplored. A research methodology is instituted, and statistics of the findings are presented, as well as a synthesis of the main methods identified. It is concluded that the optimization of PD is still an open question and that the use of AM in RC still lacks comprehensive studies. Furthermore, the reviewed works evidenced a deficiency in reproducibility and comparability of the results obtained.*

Keywords. *Debt Collection, Machine Learning, Reinforcement Learning, Systematic Literature Review*

Resumo. *O mercado de Recuperação de Créditos (RC) apresenta um vasto potencial para aprimoramento seus processos através da implementação de técnicas de Aprendizado de Máquina (AM), devido à abundância de dados agregados por instituições de cobrança. Esta Revisão Sistemática da Literatura tem como objetivo identificar métodos aplicados ao problema específico de Precificação de Dívidas (PD), em especial Aprendizado por Reforço, bem como destacar as oportunidades de pesquisa ainda não exploradas. É estabelecida uma metodologia de busca e são apresentadas estatísticas dos resultados encontrados, além de uma síntese dos principais métodos identificados. Conclui-se que a otimização da PD é uma questão em aberto e que o uso de AM em RC ainda carece de estudos abrangentes. Além disso, observou-se uma falta de reprodutibilidade e comparabilidade entre os resultados obtidos nos trabalhos analisados.*

Palavras-chave. *Recuperação de Crédito, Aprendizado de Máquina, Aprendizado por Reforço, Revisão Sistemática da Literatura*

1. Introdução

De acordo com dados de [Barbosa 2023], há cerca de 65 milhões de brasileiros inadimplentes, totalizando R\$ 248 bilhões em dívidas. Muitas empresas credoras não possuem a infraestrutura necessária para contatar e negociar pendências com seus clientes e, por isso, vendem o direito de cobrança e recebimento dessas dívidas para empresas especializadas, denominadas instituições de cobrança ou recuperação de crédito. Essas instituições, por sua vez, agregam informações de diferentes fontes sobre um mesmo cliente ou dívida

e conseguem otimizar o processo de cobrança focando em casos com maior probabilidade de retorno. A grande quantidade de dados disponíveis para estas instituições torna o Aprendizado de Máquina (AM) uma ferramenta promissora para a otimização de seus processos.

O ciclo de vida do crédito, conforme demonstrado pela Figura 1, é constituído por diversas fases que podem incluir clientes e instituições interessadas. A princípio, um cliente submete uma aplicação para crédito em uma instituição credora. Essa instituição realiza a avaliação de crédito que envolve, entre outros, a avaliação do risco de inadimplência. Após a aprovação do cliente pela instituição, esta procede com a precificação do crédito, avaliando a taxa de juros a ser aplicada. Caso o cliente não efetue o pagamento no prazo preestabelecido, a instituição pode contatá-lo para realizar as primeiras cobranças. Se a instituição de crédito considerar inviável a cobrança, ela pode vender a dívida em um grupo de créditos com características similares, conhecidos como carteiras ou portfólios. Instituições especializadas em recuperação de crédito podem adquirir essas carteiras por meio de leilões, avaliando sua rentabilidade e oferecendo lances. Após a compra, a instituição de recuperação deve decidir o desconto a ser dado em cada dívida, a fim de aumentar a taxa de conversão em pagamentos. Além disso, deve também escolher quais clientes contatar, uma vez que o número de recebíveis pode extrapolar o viável pela instituição e as informações de contato podem estar incorretas. Após o pagamento da dívida, é importante para as instituições envolvidas avaliar a contribuição de cada decisão tomada ao longo do ciclo e reconhecer o mérito de cada uma, a fim de otimizar seus processos. Após o pagamento da dívida, é fundamental que as instituições envolvidas avaliem cada decisão tomada ao longo do ciclo e reconheçam oportunidades de otimização em seus processos. O problema de recuperação de crédito abordado neste estudo consiste, principalmente, nas etapas de precificação da dívida e decisão de acionamento, podendo ser auxiliado por soluções em precificação de carteiras e reconhecimento de mérito.

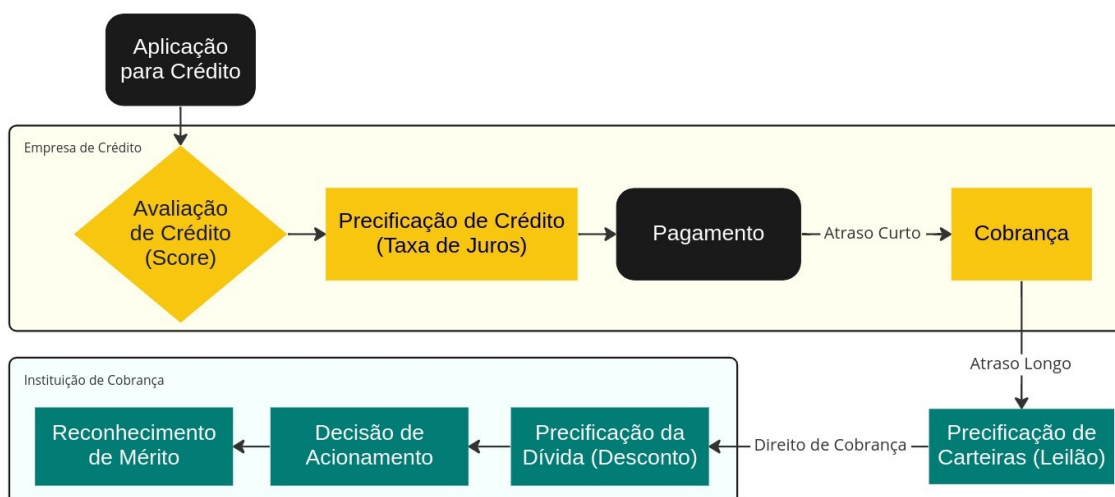


Figura 1. Ciclo de vida típico do crédito no caso de não-pagamento nos prazos pré-estabelecidos. As etapas de Aplicação para Crédito e Pagamento são executadas pelo cliente, enquanto as demais são executadas por empresas de crédito ou instituições de cobrança. Este trabalho foca em soluções aplicáveis ao passo de precificação da dívida, ou seja, a escolha do desconto sobre uma dívida.

Cada etapa do ciclo de vida do crédito pode ser otimizada de forma independente. A etapa de Avaliação de Crédito, ou Score, por exemplo, é amplamente estudada no contexto de aprendizado supervisionado devido a aplicabilidade de métodos estabelecidos sobre os dados desse problema [Dastile et al. 2020]. No entanto, a natureza sequencial e interativa do ciclo de crédito sugere que métodos como o Aprendizado por Reforço (AR), capazes modelar e se beneficiar dessas características, sejam mais adequados para a tomada de decisão. Por exemplo, [Oliveira et al. 2021] propõe uma forma de otimizar a etapa de decisão de acionamento com AR *Offline*, uma variação de AR capaz de se aproveitar de grandes bases de dados históricas para parametrização prévia e posterior implantação em um ciclo de otimização *online*, como parte do sistema de uma empresa parceira.

Esta revisão é parte de uma pesquisa que busca testar duas principais hipóteses. A primeira afirma que é possível utilizar grandes bases históricas para otimizar o processo de recuperação de crédito através da personalização da precificação de dívidas. A segunda afirma que a relação de longo prazo de um cliente com seu crédito e instituições de cobrança oferece informações úteis para esta otimização. O objetivo do presente trabalho é, portanto, identificar o estado da arte de aplicações de AM sobre o problema de recuperação de crédito e, mais especificamente, validar o uso de AR e AR *Offline* para a otimização da escolha de descontos.

Este trabalho está organizando da seguinte maneira: na Seção 2 é apresentada a metodologia da revisão, de forma a garantir sua reprodutibilidade; na Seção 3 são apresentadas estatísticas e respostas extraídas através da condução da revisão; na Seção 4 é feita uma análise dos métodos identificados e principais abordagens a problemas correlatos e; na Seção 5 é apresentada uma síntese dos resultados e um levantamento de oportunidades de pesquisa para trabalhos futuros.

2. Metodologia da Pesquisa

Nesta seção é apresentado o planejamento da revisão sistemática, que foi executada com apoio do software Parsifal¹. Esse planejamento tem como objetivo garantir seu rigor e reprodutibilidade, além de permitir o levantamento abrangente de estudos relevantes e a identificação e análise de métodos importantes. Esta seção está dividida entre a definição das Questões de Pesquisa, levantamento das Palavras-Chave e Sinônimos, especificação das Bases de Busca, Critérios de Inclusão e Exclusão de estudos, e Questões de Qualidade para classificação de trabalhos aceitos.

2.1. Questões de Pesquisa

Para a realização desta revisão, foram definidas quatro Questões de Pesquisa (QP) baseadas nas hipóteses gerais do projeto, apresentadas na Seção 1. Essas questões são listadas a seguir:

- QP.1 Quais etapas do processo de recuperação de crédito foram abordadas com aprendizado de máquina?
- QP.2 Quais as principais técnicas de aprendizado de máquina aplicadas ao problema abordado?

¹Disponível em <https://parsif.al>

- QP.3 Como características sequenciais dos processos de decisão foram modeladas?
QP.4 Foram propostos *benchmarks* ou *datasets* públicos para o problema? Que métricas são avaliadas?

2.2. Palavras-Chave e Sinônimos

Foram definidas palavras-chave (K) relacionadas ao contexto de recuperação de créditos e às intervenções baseadas em aprendizado de máquina e aprendizado por reforço. Essas palavras-chave têm o objetivo de serem abrangentes a fim de montar um panorama da atividade científica sobre cada subproblema relacionado à recuperação de créditos.

- K.1 {Recuperação de Crédito (Negociação de Dívidas, Debt Collection, Consumer Credit) }
K.2 {Aprendizado de Máquina (Machine Learning, Aprendizagem de Máquina)}
K.3 {Aprendizado por Reforço (Reinforcement Learning, Aprendizagem por Reforço)}

Entre parênteses, são especificados sinônimos ou termos utilizados por literatura correlata. A cadeia (S) lógica de busca foi definida como $S = [K.1] \wedge ([K.2] \vee [K.3])$. A cadeia resultante é apresentada a seguir:

$$S = [“recuperação de crédito” \vee “negociação de dívidas” \vee “debt collection” \vee “consumer credit”] \wedge [[“machine learning” \vee “aprendizado de máquina” \vee “aprendizagem de máquina”] \vee [“reinforcement learning” \vee “aprendizado por reforço” \vee “aprendizagem por reforço”]]$$

2.3. Bases de Busca

A seleção de repositórios para esta revisão se baseou na relação com o tema da pesquisa, reprodutibilidade das buscas e o número de trabalhos científicos indexados. As bases de busca escolhidas foram: ACM Digital Library, IEEE Digital Library, ISI Web of Science, Science@Direct, Scopus e Springer Link. Em todas as bases consideradas, a cadeia lógica de busca S foi aplicada nos campos de título, resumo, palavras chaves e texto corrido dos artigos.

2.4. Critérios de Inclusão e Exclusão

Todos os resultados da busca foram avaliados e filtrados segundo critérios objetivos de inclusão e exclusão, a fim de garantir sua relevância ao tema e às perguntas a serem respondidas. Trabalhos poderão ser aceitos caso satisfaçam pelo menos um dos seguintes critérios de inclusão:

- I.1 O artigo científico aborda o problema de precificação de dívidas;
- I.2 A solução apresentada considera ou modela características sequenciais do problema abordado;
- I.3 A solução se beneficia do uso de grandes bases de dados ou propõe método para sintetização dos mesmos;
- I.4 O artigo propõe metodologia ou técnica relevante para as questões de pesquisa.

Resultados que não se encaixem em nenhum dos critérios mencionados são automaticamente rejeitados. Além disso, resultados também são rejeitados caso satisfaçam algum dos seguintes critérios de exclusão:

Bases de Busca	Estudos		Duplicatas		Rejeitados		Aceitos	
	Qtde		Qtde	%	Qtde	%	Qtde	%
ACM Digital Library	18		1	5,5	16	88,9	1	5,5
IEEE Digital Library	12		3	25	9	75	0	0
ISI Web of Science	38		13	34,2	24	63,1	1	2,6
Science@Direct	63		7	11,1	54	85,7	2	3,1
Scopus	21		0	0	17	80,9	4	19,1
Springer Link	129		3	2,3	126	97,7	0	0
Total	281		27	9,6	246	87,5	8	2,8

Tabela 1. Número de estudos por categoria, em cada base de busca.

- E.1 Documento não se refere ao problema de recuperação de crédito;
- E.2 Solução não é automatizada com aprendizado de máquina;
- E.3 Documento não é um artigo científico publicado em anais de conferência ou periódicos com revisão por pares;
- E.4 Documento não está nos idiomas inglês ou português.

Portanto, trabalhos aceitos devem, ao mesmo tempo, satisfazer algum critério de inclusão e nenhum critério de exclusão.

$$I.1 \vee [[I.2 \vee I.3 \vee I.4] \wedge \neg[E.1 \vee E.2 \vee E.3 \vee E.4 \vee E.5 \vee E.6]]$$

2.5. Questões de Qualidade

As Questões de Qualidade (QQ) foram definidas a fim de quantificar a relevância dos resultados da busca para as questões de pesquisa. As questões a seguir são avaliadas para todos os trabalhos aceitos na etapa anterior, definida na Subseção 2.4:

- QQ.1 O trabalho aborda o problema de precificação de dívidas?
- QQ.2 A solução apresentada considera características sequenciais do problema abordado?
- QQ.3 A solução apresentada é capaz de se adaptar após a parametrização inicial (treinamento), ou seja, durante a execução?
- QQ.4 O trabalho apresenta uma técnica aplicável ao problema de precificação de crédito?
- QQ.5 Os resultados são reprodutíveis e comparáveis com trabalhos relacionados?
- QQ.6 Os experimentos são rigorosos e os resultados são abrangentes e convincentes?
- QQ.7 Proposta é avaliada em ambiente relevante (e.g. em produção)?

As respostas possíveis para cada pergunta são “Sim”, “Não” ou “Parcialmente”, correspondendo a pesos 1, 0 e 0.5, respectivamente. Para cada estudo avaliado, os pesos dessas respostas são somados em uma nota e apresentados na coluna “Total” da Tabela 2. Neste trabalho, a nota atribuída a cada estudo é utilizada como métrica de importância e prioridade durante sua avaliação e levantamento de oportunidades de pesquisa.

3. Condução e Resultados

Para a etapa de condução da revisão sistemática foi aplicada a cadeia lógica de busca S, apresentada na Subseção 2.2, sobre as bases listadas na Subseção 2.3. Uma exceção,

no entanto, foi a busca executada sobre a base Science@Direct, que precisou ser alterada para se adequar a limitação de, no máximo, oito conectores lógicos por cadeia. Essa alteração consistiu da aplicação da regra lógica distributiva da conjunção sobre a disjunção [Rosen 2018] na cadeia original, a busca independente das conjunções e a união posterior (i.e. disjunção) dos resultados.

A Tabela 1 apresenta o número de trabalhos por categoria da revisão. A ordem de análise dos resultados da busca foi a seguinte: Scopus, ACM Digital Library, IEEE Digital Library, Springer Link, Science@Direct e ISI Web of Science. Como diversas bases podem indexar um mesmo estudo, resultados importados que já haviam sido analisados foram considerados duplicados, o que explica o número de duplicatas em cada base de busca. Além disso, a busca executada na base Springer Link retornou um número desproporcional de resultados pois não foi estrita o suficiente, devido à plataforma aplicar a técnica de *stemming* nas palavras da cadeia lógica. Como resultado, muitos estudos não estavam diretamente relacionados ao problema de recuperação de créditos ou abordavam questões fora do escopo desta pesquisa, sendo, portanto, rejeitados.

Em resumo, todos os artigos aceitos satisfazem os critérios de inclusão I.2 e I.3 em algum nível. No entanto, nenhum resultado da busca satisfaz o critério I.1, o que, dada a abrangência das palavras-chave consideradas, sugere que o problema de precificação de dívidas é um problema ainda não abordado na literatura revisada por pares no contexto de aprendizado de máquina.

No total, foram aceitos oito estudos para análise:

- P.1 [Abe et al. 2010]
- P.2 [Kennedy et al. 2013]
- P.3 [Herasymovych et al. 2019]
- P.4 [Yang et al. 2020]
- P.5 [Park et al. 2021]
- P.6 [Kim et al. 2021]
- P.7 [Kuzmin et al. 2022]
- P.8 [Khraishi and Okhrati 2022]

A Tabela 2 apresenta as respostas para as perguntas de qualidade de cada estudo. Para a QQ.1, as respostas foram “Parcialmente” para estudos cujas abordagens poderiam ser adaptadas para o problema de precificação de dívidas, já que nenhum estudo abordou diretamente essa questão. Na QQ.2, as respostas foram “Parcialmente” para estudos que consideraram as características sequenciais apenas na seleção de variáveis, mas não na técnica utilizada para solucionar o problema. Na QQ.4, as respostas foram “Parcialmente” para estudos que apresentaram uma técnica útil de forma indireta, podendo requerer uma reformulação do problema para ser aplicada à precificação de dívidas. As demais respostas seguem a interpretação convencional de cada pergunta.

A Tabela 3 mostra, de forma resumida, as respostas para as questões de pesquisa para cada estudo. Essas respostas são discutidas de forma mais detalhada na Seção 4.

4. Métodos Identificados

Nesta seção são apresentados os métodos identificados nos oito artigos científicos aceitos durante a condução da revisão, em relação às questões de pesquisa definidos durante seu planejamento.

	QQ.1	QQ.2	QQ.3	QQ.4	QQ.5	QQ.6	QQ.7	Total
P.1	Não	Sim	Sim	Sim	Sim	Sim	Sim	6.0
P.2	Não	Parc.	Não	Não	Parc.	Sim	Não	2.0
P.3	Parc.	Sim	Sim	Parc.	Não	Parc.	Sim	4.5
P.4	Não	Sim	Sim	Parc.	Parc.	Sim	Parc.	4.5
P.5	Não	Parc.	Não	Sim	Não	Sim	Não	2.5
P.6	Não	Sim	Não	Sim	Parc.	Sim	Não	3.5
P.7	Parc.	Sim	Sim	Parc.	Não	Não	Não	3.0
P.8	Parc.	Sim	Sim	Sim	Sim	Sim	Não	5.5

Tabela 2. Respostas às Questões de Qualidade. A coluna “Total” se refere à soma dos valores atribuídos a cada resposta, definido na Subseção 2.5

	QP.1	QP.2	QP.3	QP.4
P.1	Acionamento	AR	Tempo da dívida, canal	Público, Online; Lift
P.2	Score	RL	Prob. de pagamento	Público; Acurácia
P.3	Taxa de Juros	AR	Concessão semanal	Privado, Online; Lift
P.4	Acionamento	AR	Tempo da dívida, canal	Privado, Online; Lift
P.5	Score	RAG	Histórico de dívidas	Privado; Consistência
P.6	Score	AEV	Faturas mensais de cartão	Privado; F1 Score
P.7	Acionamento	AR Offline	Tempo da dívida, canal	Privado; DT
P.8	Taxa de Juros	AR Offline	Aplicações para crédito	Público, Sintético; MD

Tabela 3. Resumo das respostas às Questões de Pesquisa. Alguns métodos identificados são: Aprendizado por Reforço (AR), Regressão Logística (RL), Redes Adversárias Generativas (RAG), Autoencoders Variacionais (AEV), Diferença Temporal (DT) e Método Direto (MD).

O primeiro método, proposto por [Abe et al. 2010], utiliza um processo de decisão de Markov restrito e aprendido por reforço para escolher o melhor canal de comunicação para contactar cada cliente, levando em consideração o tempo decorrido desde o registro da dívida. A abordagem foi testada em datasets públicos e em um ambiente online, para otimização de cobrança de impostos no Departamento de Tributação e Finanças do Estado de Nova York. Seus resultados comparam a recompensa total obtida e o *lift* das métricas de conversão em relação às políticas de acionamento originais do sistema.

O estudo de [Yang et al. 2020], por sua vez, utiliza aprendizado por reforço para otimizar a estratégia de acionamento e cobrança de dívidas em plataformas de empréstimos. Seus experimentos utilizam um dataset privado e seus testes em ambiente online não são bem descritos. Em suas conclusões, os autores destacam a importância da personalização na cobrança de dívidas e sugerem que as plataformas de empréstimos utilizem ações mais cautelosas para cobrança, valorizando a privacidade de seus clientes.

[Kuzmin et al. 2022] também aborda o problema de acionamento, mas sua abordagem é capaz de absorver e otimizar políticas em grandes bases de dados com AR Offline. No entanto, sua modelagem não é bem definida por questões de confidencialidade e o dataset utilizado para parametrização e avaliação do método também é privado. O método desenvolvido é avaliado contra outros algoritmos de aprendizado de máquina, e é demonstrada a vantagem do aprendizado por reforço profundo em termos de precisão e eficiência. A métrica de avaliação utilizado é a perda de recompensa com Diferença Temporal (DT) [Sutton 1988].

Já [Herasymovych et al. 2019] utiliza AR em seu trabalho na etapa de precificação de créditos, escolhendo a taxa de juros a ser aplicada a cada concessão avaliada. O algoritmo deve identificar a taxa de aceitação dos clientes, aproximando-a da taxa de concessão definida semanalmente pela empresa em que é aplicado. Embora o dataset utilizado para a parametrização do método seja privado, a solução testada online é bem descrita e o lucro obtido é comparado com o lucro das políticas originais da empresa.

[Khraishi and Okhrati 2022] também otimiza a escolha da taxa de juros em seu estudo, mas com AR Offline. Sua abordagem faz uso de grandes bases históricas de aplicações a crédito para otimizar a taxa de juros e minimizar o não-pagamento de empréstimos ao longo da vida de um mesmo cliente, sem estabelecer suposições sobre a forma funcional da demanda. O estudo também apresenta um dataset público criado sinteticamente, onde o mesmo avalia suas soluções através do Método Direto (MD) [Dudík et al. 2014].

Muitos trabalhos retornados pelas buscas são relacionados à etapa de Avaliação de Crédito, ou Score. Um deles, apresentado por [Kennedy et al. 2013], utiliza Regressão Logística (RL) para determinar a probabilidade de pagamento de empréstimos em janelas fixas de tempo. O objetivo é caracterizar o comportamento de pagamento de clientes e o risco de novos empréstimos. O dataset utilizado é baseado em dados públicos do mercado irlandês e a acurácia da solução proposta é quantificada. De acordo com os autores, um período de desempenho de 12 meses é o mais adequado para realizar previsões com maior precisão entre os períodos considerados.

Outros estudos, como os de [Park et al. 2021] e [Kim et al. 2021], também abor-

dam o Score, mas sob perspectivas diferentes. Em [Park et al. 2021] é proposto um método para sintetização de datasets com Redes Adversárias Generativas (RAG), garantindo a consistência de métricas e dependências estatísticas entre o dataset original, que é privado, e o dataset sintetizado. Essa abordagem permite que modelos de aprendizado de máquina sejam desenvolvidos e comparados diretamente, atacando uma limitação comum nesta área de pesquisa. Por outro lado, [Kim et al. 2021] utiliza Autoencoders Variacionais (AEV) para gerar representações semânticas de clientes com base em seus pagamentos de faturas de cartão de crédito. Isso possibilita determinar a probabilidade de inadimplência em uma sequência de faturas.

5. Conclusões

Foi constatado que o Aprendizado por Reforço já é uma técnica utilizada para otimizar a tomada de decisões em relação ao acionamento de clientes e à precificação de crédito, ambos de grande importância no ciclo de vida do crédito. Ademais, os avanços recentes nesta técnica podem contribuir ainda mais para o aprimoramento do setor.

Embora a literatura revisada não tenha explorado especificamente a precificação de dívidas, as abordagens identificadas podem ser adaptadas para solucionar o problema em questão. A modelagem proposta por [Kennedy et al. 2013], por exemplo, pode auxiliar na definição de uma função de recompensa baseada em probabilidade de pagamento para agentes treinados em um conjunto de dados históricos de clientes. As métricas utilizadas por [Abe et al. 2010], [Kuzmin et al. 2022] e [Khraishi and Okhrati 2022] também podem fornecer uma valiosa ferramenta para avaliar o desempenho de futuras soluções. Além disso, a incorporação de técnicas de aprendizado de representações, como sugerido por [Kim et al. 2021] pode tornar escolhas feitas por algoritmos mais interpretáveis e simplificar o processo de tomada de decisões em diferentes etapas. Por fim, destaca-se a ausência de benchmarks estabelecidos para avaliar e comparar as soluções propostas na literatura. Nesse sentido, a sintetização de conjuntos de dados, conforme sugerido por [Park et al. 2021], pode ser uma alternativa interessante para a criação de um ponto de comparação comum entre os estudos nesta área, considerando a sensibilidade dos dados envolvidos.

Referências

- Abe, N., Melville, P., Pendus, C., Reddy, C. K., Jensen, D. L., Thomas, V. P., Bennett, J. J., Anderson, G. F., Cooley, B. R., Kowalczyk, M., et al. (2010). Optimizing debt collections using constrained reinforcement learning. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 75–84.
- Barbosa, M. (2023). Inadimplência tem pequena queda e atinge 65,06 milhões de brasileiros. In de Dirigentes Lojistas (CNDL) e Serviço de Proteção ao Crédito, C. N., editor, *Confederação Nacional de Dirigentes Lojistas (CNDL) e Serviço de Proteção ao Crédito*. Confederação Nacional de Dirigentes Lojistas (CNDL) e Serviço de Proteção ao Crédito. Disponível em <https://cndl.org.br/varejosa/inadimplencia-tem-pequena-queda-e-atinge-6506-milhoes-de-brasileiros/>. Acessado em 18/02/2023.
- Dastile, X., Celik, T., and Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91:106263.

- Dudík, M., Erhan, D., Langford, J., and Li, L. (2014). Doubly robust policy evaluation and optimization. *Statistical Science*, 29(4).
- Herasymovych, M., Märka, K., and Lukason, O. (2019). Using reinforcement learning to optimize the acceptance threshold of a credit scoring model. *Applied Soft Computing*, 84:105697.
- Kennedy, K., Mac Namee, B., Delany, S. J., O’Sullivan, M., and Watson, N. (2013). A window of opportunity: Assessing behavioural scoring. *Expert Systems with Applications*, 40(4):1372–1380.
- Khraishi, R. and Okhrati, R. (2022). Offline deep reinforcement learning for dynamic pricing of consumer credit. In *Proceedings of the Third ACM International Conference on AI in Finance*, pages 325–333.
- Kim, Y., Cho, Y., Lee, H., and Moon, I.-C. (2021). Predict sequential credit card delinquency with vade-seq2seq. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1159–1164. IEEE.
- Kuzmin, G., Panov, A. I., Razvorotnev, I., and Rezyapkin, V. (2022). Application of deep reinforcement learning methods in debt collection. In *Proceedings of the Fifth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’21)*, pages 66–77. Springer.
- Oliveira, B., Melo, L., Martins, L., Brandao, B., Soares, D., and Lima, T. (2021). Pulserl: Enabling offline reinforcement learning for digital marketing systems via conservative q-learning. In *2nd Offline Reinforcement Learning Workshop at the 35th Conference on Neural Information Processing*.
- Park, N., Gu, Y. H., and Yoo, S. J. (2021). Synthesizing individual consumers’ credit historical data using generative adversarial networks. *Applied Sciences*, 11(3):1126.
- Rosen, K. H. (2018). *Discrete Mathematics and Its Applications*. McGraw-Hill Education.
- Sutton, R. (1988). Learning to predict by the method of temporal differences. *Machine Learning*, 3:9–44.
- Yang, C., Lu, T., Li, B., and Lu, X. (2020). Personalizing debt collections: Combining reinforcement learning and field experiment. *International Conference on Information Systems, ICIS 2020 - Making Digital Inclusive: Blending the Local and the Global*.