UNIVERSIDADE FEDERAL DE JUIZ DE FORA
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

JUDY THULER DE ALMEIDA

VIÉS ALGORÍTMICO NAS DECISÕES ENVOLVENDO *MACHINE LEARNING* EM NEGÓCIOS SOCIAIS LIGADOS AO MICROCRÉDITO

JUIZ DE FORA
2023

JUDY THULER DE ALMEIDA

VIÉS ALGORÍTMICO NAS DECISÕES ENVOLVENDO *MACHINE LEARNING* EM NEGÓCIOS SOCIAIS LIGADOS AO MICROCRÉDITO

Trabalho de Conclusão de Curso apresentado à Faculdade de Engenharia da Universidade Federal de Juiz de Fora, como requisito parcial para a obtenção do título de Engenheiro de Produção.

Orientador: D. Sc., Mariana Paes da Fonseca
Co-Orientador: D. Sc, Bruno Milanez

JUIZ DE FORA

2023

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Almeida, Judy Thuler de.
VIÉS ALGORÍTMICO NAS DECISÕES ENVOLVENDO MACHINE LEARNING EM NEGÓCIOS SOCIAIS LIGADOS AO MICROCRÉDITO / Judy Thuler de Almeida. -- 2023.
72 f.

Orientadora: Mariana Paes da Fonseca
Coorientador: Bruno Milanez
Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, Faculdade de Engenharia, 2023.

1. Machine Learning. 2. Microcrédito. 3. Viés algorítmico. I. Fonseca, Mariana Paes da , orient. II. Milanez, Bruno, coorient. III. Título.

JUDY THULER DE ALMEIDA

VIÉS ALGORÍTMICO NAS DECISÕES ENVOLVENDO *MACHINE LEARNING* EM NEGÓCIOS SOCIAIS LIGADOS AO MICROCRÉDITO

Trabalho de Conclusão de Curso apresentado à Faculdade de Engenharia da Universidade Federal de Juiz de Fora, como requisito parcial para a obtenção do título de Engenheiro de Produção.

Aprovada em 23 de junho de 2023.

BANCA EXAMINADORA

D. Sc., Mariana Paes da Fonseca (Orientadora)
Universidade Federal de Juiz de Fora

D. Sc., Bruno Milanez (Coorientador)
Universidade Federal de Juiz de Fora

D. Sc., Marcos Vinicius Rodrigues
Universidade Federal de Juiz de Fora

AGRADECIMENTOS

Agradeço a todos que fizeram parte da minha graduação. Um agradecimento especial aos meus pais, Maria Luiza e Edson, que foram minha base ao longo de toda a minha vida, que sempre apoiaram e suportaram minhas escolhas. Nada disso seria possível sem o apoio de vocês! Agradeço também a toda minha família, que comemorou minhas vitórias e me apoiou nos momentos difíceis.

Ao meu namorado e companheiro de faculdade, Pedro Henrique, que sempre me suportou e foi meu porto seguro. Os últimos 4 anos não seriam os mesmo sem você!

As minhas amigas de escola, Maria Eduarda e Lara que, mesmo através da distância, se fizeram presentes durante os últimos anos.

Aos meus amigos de graduação, com destaque especial para Luciano Bittencourt, Luca Salles, Matheus Santos, Laís Gonçalves, Lorrayne Duarte e Gustavo Gayer.

A UFJF, que foi palco de tanto aprendizado e crescimento, e todos os envolvidos para o seu funcionamento. Durante o período da minha graduação tive a oportunidade de contar com o apoio de diferentes segmentos, que agregaram na minha formação acadêmica. Sou extremamente grata ao desenvolvimento adquirido nas passagens à SEEPRO, CRITT E ENACTUS UFJF. Um agradecimento especial à gestão 2019. 1 da SEEPRO, na qual fiz amigos que levarei para toda a vida.

A minha orientadora, Mariana, sou grata por ter me suportado durante todo o meu TCC. Obrigada por me mostrar os melhores caminhos durante os últimos três períodos. Obrigada também por disseminar tanto conhecimento ao longo das suas disciplinas.

Ao meu coorientador, Bruno, obrigada por todo apoio e sugestões. Obrigada por concordar em fazer parte desse estudo, e fornecer sempre uma visão mais humana da engenharia. Suas aulas sempre foram inspiração para mim.

RESUMO

O empreendedorismo social é uma vertente econômica que vem crescendo ao longo dos anos. Visando suprir necessidades que nem sempre são atendidas pelo mercado tradicional, empreendedores sociais buscam solucionar problemas e ajudar a comunidade. Nesse campo, surge o microcrédito, que é uma forma de conceder empréstimos para a população que não consegue recursos em bancos convencionais. Para auxiliar no modelo de negócio do microcrédito, podem ser utilizadas as ferramentas de *Machine Learning*, que tratam os dados e fornecem saídas de grande valia para usuários e fornecedores. No entanto, a tecnologia em questão pode apresentar falhas que atenuam ainda mais as lacunas sociais, como é o caso do viés algorítmico. Diante disso, através de uma revisão bibliográfica e documental, com estudo de caso da instituição Kiva.org, serão avaliadas as possíveis formas de viés algorítmico na utilização de *Machine Learning* atrelada ao microcrédito, além de propor soluções baseadas em estudos prévios. Serão explicitadas as possibilidades de viés nas ferramentas do site, como formulário de inscrição, ferramenta de busca, *tags* e recomendações. Por fim, serão resumidas as possibilidades de viés em cada método de aprendizado, além de ressaltar a importância das ferramentas de *Machine Learning* para estudos futuros.

Palavras-chave: *Machine Learning*, microcrédito, viés algorítmico.

ABSTRACT

Social entrepreneurship is an economic trend that has been growing over the years. Aimed at meeting needs that are not always addressed by the traditional market, social entrepreneurs seek to solve problems and help the community. In this field, microcredit emerges as a way to provide loans to people who cannot access conventional banks. To support the microcredit business model, Machine Learning tools can be used, which process data and provide valuable outputs for users and providers. However, the technology in question can have flaws that further exacerbate social gaps, such as algorithmic bias. Therefore, through a literature and documentary review, with a case study of the institution Kiva.org, possible forms of algorithmic bias in the use of Machine Learning linked to microcredit will be evaluated, as well as proposing solutions based on previous studies. The possibilities of bias in the website's tools, such as the application form, search tool, tags, and recommendations, will be explained. Finally, the possibilities of bias in each learning method will be summarized, highlighting the importance of Machine Learning tools for future studies.

Keywords: Machine Learning, microcredit, algorithmic bias.

LISTA DE FIGURAS

Figura 1 - Os 17 ODS 15
Figura 2 - Teor multidisciplinar do empreendedorismo social 22
Figura 3 - Tipos de Aprendizado de *Machine Learning* 31
Figura 4 - Problema de Liberação de Crédito 33
Figura 5 – Modelo de Regressão 33
Figura 6 – Dados do Modelo de Regressão 34
Figura 7 – Matriz de Valores para a Resposta Predita 40
Figura 8 – Categorias de cadastro Kiva: condenação e falência 45
Figura 9 – Categorias do Site Kiva.org 47
Figura 10 - Categorias de cadastro Kiva: valor do empréstimo 48
Figura 11 – Categoria de tendência 48
Figura 12 – Recomendações 50
Figura 13 – Recomendações do site Kiva 50
Figura 14 – Total de empréstimos do site Kiva.org por região 52
Figura 15 – Total de recomendações por região 52

LISTA DE TABELAS

Tabela 1- Ranking de IDH, países selecionados 14
Tabela 2 – Resumo Vieses..... 54
Tabela 3 – Resumo: Métodos de análise de saídas..... 55
Tabela 4 – Quadro resumo da classificação dos vieses 56

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ABEPRO – ASSOCIAÇÃO BRASILEIRA DE ENGENHARIA DE PRODUÇÃO

BOP - BOTTOM OF THE PYRAMID (BASE DA PIRÂMIDE)

DS – DATA SCIENCE

ENACTUS - ENTREPRENEURIAL ACTION US

IDH - ÍNDICE DE DESENVOLVIMENTO HUMANO

ML – MACHINE LEARNING

ODS - OBJETIVOS DE DESENVOLVIMENTO SUSTENTÁVEL

ONGS - ORGANIZAÇÕES NÃO GOVERNAMENTAIS

ONU - ORGANIZAÇÃO DAS NAÇÕES UNIDAS

UNDP - UNITED NATIONS DEVELOPMENT PROGRAMME

UNICEF – FUNDO DAS NAÇÕES UNIDAS PARA A INFÂNCIA

SUMÁRIO

1. INTRODUÇÃO 12
1.1 CONSIDERAÇÕES INICIAIS 12
1.2 JUSTIFICATIVA 13
1.3 ESCOPO DO TRABALHO 18
1.4 ELABORAÇÃO DOS OBJETIVOS 18
1.5 DEFINIÇÃO DA METODOLOGIA 19
1.6 ESTRUTURA DO TRABALHO 19
2. EMPREENDEDORISMO SOCIAL, MICROCRÉDITO DATA SCIENCE E CONCEITOS CORRELATOS 21
2.1 EMPREENDEDORISMO SOCIAL 21
2.1.1 DEFINIÇÃO DE EMPREENDEDORISMO SOCIAL 21
2.1.2 EMPREENDEDORISMO SOCIAL NO MICROCRÉDITO 23
2.2 DATA SCIENCE 27
2.2.1 CONCEITO 27
2.2.2 CAMPOS DE APLICAÇÃO 29
2.2.3 *MACHINE LEARNING* 30
2.2.4 MODELOS DE ANÁLISE DAS SAÍDAS 32
3. VIÉS ALGORÍTMICO E O EMPREENDEDORISMO SOCIAL COM ENFOQUE NO MICROCRÉDITO 34
3.1 VIÉS ALGORÍTMICO E SEUS RISCOS 34
3.2 *MACHINE LEARNING*: DIFERENTES FORMAS DE APRENDIZADO E O VIÉS ALGORÍTMICO 38
3.2.1 APRENDIZADO SUPERVISIONADO 38
3.2.2 APRENDIZADO NÃO SUPERVISIONADO 40
3.2.3 APRENDIZADO SEMI SUPERVISIONADO 42
3.2.4 APRENDIZADO POR REFORÇO 42
3.3 POSSIBILIDADES E CONSEQUÊNCIAS DO VIÉS ALGORÍTMICO NO MICROCRÉDITO 44
3.3.1 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE CLASSIFICAÇÃO 45
3.3.2 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE REGRESSÃO 47
3.3.3 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE AGRUPAMENTOS 49
3.3.4 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE ASSOCIAÇÃO 50
4. CONCLUSÕES 54
REFERÊNCIAS 60
ANEXO A – TERMO DE AUTENTICIDADE 71
ANEXO B – SOLICITAÇÃO DE ACESSO AO CÓDIGO KIVA 72

1. INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

O presente trabalho tem como foco estudar as aplicações das técnicas de *Machine Learning (ML)* em empreendimentos sociais com atuação em microcrédito, evidenciando um possível viés algorítmico e suas formas de mitigação. O objetivo do estudo é fazer uma revisão bibliográfica que correlaciona conceitos de *Data Science (DS)*, empreendedorismo social e viés algorítmico com conceitos de microcrédito. Serão avaliados estudos sobre os temas supracitados, para que seja feita uma relação entre eles, mostrando como os mesmos se afetam positiva e negativamente. Para isso, serão considerados quatro tipos de análise de *Machine Learning*, verificando a possibilidade de viés em cada um deles, e buscando por soluções que se apliquem no contexto do microcrédito. Os tipos de análise considerados foram aprendizado supervisionado, aprendizado não-supervisionado, aprendizado semi-supervisionado e aprendizado por reforço.

Entrando no empreendedorismo social, é feita a apresentação de sua definição, citando as principais diferenças comparadas ao empreendedorismo tradicional. O termo empreender está ligado ao conceito de iniciativa, resolução de problemas, estímulo econômico e a busca por novas maneiras de solucionar problemas antigos (SCHAEFER, 2018). O empreendedorismo social é um ramo do mercado que é largamente conhecido como fonte de progresso e harmonia social, promovendo a inclusão, sendo fonte de empregos, criando valor social e auxiliando na luta contra a pobreza (PEREIRA, 2012). Para Dolabela (2008), empreender não é algo com base apenas na carreira e no mercado, mas sim que leva em consideração os valores do empreendedor, em sua visão de mundo e as relações sociais da sociedade no qual ele está inserido, o que ressalta o viés voltado para o surgimento do empreendedorismo social.

Na esfera econômica do empreendedorismo social, temos ainda o conceito de microcrédito, que surge junto com a imagem de Muhammad Yunus, baseado no empréstimo de pequenas quantias para a população carente, de maneira adequada e sustentável (FEITOSA, 2022). A geração de renda através do empréstimo de pequenas quantias, os microempréstimos, é uma forma de prover serviços financeiros para a população carente, fazendo-a capaz de gerar o autoemprego e a inserção na economia local, sendo um grande artifício para a redução da pobreza e das desigualdades sociais ao redor do mundo (LATIFEE, 2003).

Paralelamente, temos as técnicas de Data Science, como *Machine Learning* (ML), que facilitam ainda mais a vida humana. Através de algoritmos computacionais, que solucionam

problemas por meio da segmentação e clusterização[1], ML é um modo de lidar com dados eficientemente (MAHESH, 2020). No entanto, existem evidências de que há a possibilidade de intensificação das discriminações a partir do viés algorítmico. Tal viés pode ser intrínseco ao programador, que pode espelhar no código suas preferências e percepções, gerando algoritmos menos capazes de lidar com a diversidade, ou então causado pela maneira pela qual o código foi escrito, e também pelo propósito de criação do mesmo (DE CARVALHO, 2020).

Os estudos abordados no presente trabalho são apenas alguns exemplos de como o microcrédito é capaz de ajudar uma parcela da população que carece de serviços financeiros de qualidade, e provam que, assim como já foi feito por Yunus (PEREIRA; ROLO, 2012), é possível ajudar muitos doando pouco. Além disso, também exemplificam como o empreendedorismo social, o microcrédito e *Data Science* são capazes de atuarem de maneira conjunta.

Os três conceitos abordados se encontram dentro das áreas de atuação da Engenharia de Produção classificadas pela Associação Brasileira de Engenharia de Produção (ABEPRO). O conceito e as técnicas de *Machine Learning* podem se enquadrar em Inteligência Computacional, na área de Pesquisa Operacional, além de Gestão da Informação, do Conhecimento e da Tecnologia, no campo da Engenharia Organizacional. O empreendedorismo social, na vertente da redução das desigualdades sociais, atrelado ao microcrédito, abrange o lado da responsabilidade social da Engenharia de Produção, além de visar o desenvolvimento sustentável, que são subáreas da Engenharia da Sustentabilidade. Ademais, o empreendedorismo social também se enquadra na área da Engenharia Organizacional, com enfoque nas subáreas de Gestão Estratégica e Organizacional, Empreendedorismo e Gestão da Inovação (ABEPRO, 2008).

1.2 JUSTIFICATIVA

A pandemia gerou um contexto de intensificação e aceleração de uma desigualdade que ocorre há anos, e que tende a encontrar meios de reprodução e fortalecimento (BUTLER, 2020).

---

[1] A clusterização é uma maneira de identificar padrões e classificar os dados em grupos com as mesmas características (*clusters*) (PIMENTEL; DE FRANÇA; OMAR, 2003). Tal ferramenta permite entender e/ou extrair informações de um grupo de dados (OLIVEIRA,2008), que no presente trabalho tem relevância ao oferecer a possibilidade de identificação de pessoas que não possuem acesso aos serviços financeiros tradicionais.

Com o isolamento social e crise na maioria dos países, a população depende das políticas públicas adotadas pelos governantes, que podem vir a mitigar as desigualdades, ou então aprofundá-las. As desigualdades geradas são em diversas dimensões, como de gênero, geracionais, relacionadas à educação e ao mercado de trabalho (VOMMARO, 2021). No âmbito educacional, por exemplo, temos alunos em ambientes precários, sem acesso a recursos físicos e tecnológicos, às escolas e seus benefícios, além de professores sobrecarregados e gastando além de seus salários para conseguir dar aulas (NUÑES, 2020).

Como identificado por Butler (2020), a desigualdade não é uma realidade recente, e é possível apontar alguns fatores que levaram a sociedade até o cenário atual. A última década foi repleta de mudanças, rupturas e avanços causados pelo desenvolvimento dos países e suas populações. No entanto, para conseguir concretizar tais avanços, a sociedade exerceu grande pressão sobre o planeta, levando-o ao limite. A situação vivida com a Covid-19 é apenas o mais recente exemplo de que a busca pelo crescimento a qualquer custo pode desregular o ecossistema, trazendo grandes consequências para aqueles que o habitam (UNDP, 2020).

Ainda de acordo com o relatório elaborado pelo UNDP (2020), observam-se países com Índice de Desenvolvimento Humano (IDH) 0,957, enquanto outros estão na faixa de 0,394, conforme exposto na Tabela 1. Pelas definições contidas no relatório, o IDH é um índice que mede o desempenho médio de um país em três dimensões básicas: vida longa e saudável, conhecimento e condições de vida decentes.

Tabela 1- Ranking de IDH, países selecionados

País	Índice de Desenvolvimento Humano (IDH)	Posição no Ranking
Noruega	0,957	1
Suíça	0,955	2
Irlanda	0,955	3
México	0,779	76
Ucrânia	0,779	78
Brasil	0,765	84
Iraque	0,674	123
Índia	0,645	130
Uganda	0,544	160
Nigéria	0,539	161
Niger	0,394	188

Fonte*:* Adaptado de UNDP, 2020

Corroborando com os dados do UNDP, foi relatado no *World Inequality Report 2022*, realizado pelo *World Inequality Lab*, que os níveis de desigualdade hoje se assemelham aos do imperialismo ocidental do século 20, indicando que o mundo não conseguiu superar os problemas que foram herdados do século 19. Tais fatos mostram que existe uma grande desigualdade nas condições de vida entre os países, e que ainda temos um longo caminho até conseguirmos mitigar as desigualdades existentes (WORLD INEQUALITY LAB, 2022).

Entrando no contexto de desigualdades e desafios globais, a ONU surge com uma proposta de reunir os países e tentar mudar o curso da história. Dentro dessa proposta, nasce a Agenda 2030 e os 17 Objetivos de Desenvolvimento Sustentável (ODS), que incluem ações na esfera da erradicação da pobreza, promoção do bem-estar social, mudanças climáticas e preservação do meio ambiente. Dentro dos ODS, existem 169 metas que auxiliam na avaliação do sucesso do cumprimento dos mesmos (FAGUNDES; ZEIFERT; SIQUEIRA, 2020). A Figura 1 faz uma resumida exibição dos 17 ODS elaborados pela ONU.

Figura 1 - Os 17 ODS

Fonte: Adaptado de ONU, 2015

No que tange ao assunto do presente estudo, temos o ODS 10, com a seguinte descrição "Reduzir a desigualdade dentro dos países e entre eles", que é dividido em diversas metas, incluindo assuntos como inclusão política, econômica e social, adoção de políticas fiscais e salariais, aprimoramento da regulamentação de mercados e instituições financeiras e fomento ao desenvolvimento e o fluxo financeiro em Estados de maior necessidade (ONU, 2015).

Para contextualizar os assuntos ao cenário brasileiro, pode-se citar o 3º Mapa de Negócios de Impacto, elaborado pela Pipe Social, que indica um total de 34% das empresas mapeadas tendo como visão o Objetivo de Desenvolvimento Sustentável (ODS) 10, que é voltado para a redução das desigualdades sociais, enquanto 35% objetivam o ODS 8, o de promover o desenvolvimento econômico e trabalho decente (PIPE SOCIAL, 2021). No 2º Mapa de Negócios de Impacto, 30% das empresas buscavam atuar no ODS 10, enquanto 31% objetivavam o ODS 8 (PIPE SOCIAL, 2019). Tais dados mostram que os empreendedores vêm se preocupando cada vez mais com as questões sociais intrínsecas a pobreza e o acesso ao sistema financeiro, o que é uma tendência para o futuro (PIPE SOCIAL, 2021).

Com isso, é possível perceber que é de interesse global conseguir atingir um nível mínimo de desigualdade, contando com a cooperação internacional para a promoção de condições satisfatórias de bem-estar, qualidade de vida e dignidade humana, contribuindo para elos pacíficos entre as nações (KRITSKI BÁEZ, 2020).

No contexto das desigualdades, a inclusão financeira e o acesso aos serviços financeiros podem ser uma estratégia para a redução da pobreza, e o microcrédito entra nesse cenário como facilitador. Através dos microempréstimos, que não exigem grandes garantias e são menos burocráticos, a população que antes não era inserida economicamente, passa a exercer um papel duplo na economia, como produtor e consumidor, impulsionando os negócios locais e reduzindo a pobreza em escalas regionais (CERNEV; PROENÇA, 2016). Com o auxílio da tecnologia, o microcrédito pode ser escalado de maneira global, conectando investidores a credores de maneira simples e eficiente, como é o caso da instituição Kiva.org (CHOO et al., 2014)

Com o avanço e propagação da Internet e de tecnologias correlacionadas, o número de informações disponíveis se mostra cada vez maior. No entanto, a informação só se torna útil quando tratada de maneira correta, sendo capaz de trazer ativos valiosos para quem as recebe. Diante disso, surgem as técnicas de *Data Science (DS)* que, através da junção de teorias de diversas áreas, como matemática, estatística, lógica, engenharia etc., é capaz de auxiliar de maneira inteligente na tomada de decisões.

Entre as técnicas de *DS*, existem aquelas de ML, que permitem ganhos em inovação, produtividade e redução de custos, ao possibilitarem a manipulação de informações e utilização das mesmas, sendo grandes aliadas das empresas privadas e governamentais (DOS PASSOS, 2016). Entretanto, tal avanço também pode trazer consequências negativas, ao conter um viés algorítmico que transmita preferências e preconceitos de grupos sociais específicos. O comprometimento dos códigos, dado pelas saídas imparciais e comprometidas, implica em

questões éticas e discriminatórias, afetando a capacidade da ferramenta de mostrar saídas baseadas apenas nos dados inseridos, resultando em informações distorcidas da realidade que se baseiam nas opiniões pessoais do programador e dos usuários (ALVES; ANDRADE, 2021).

Trazendo o tema para o contexto brasileiro, pode-se citar novamente os estudos feitos pela Pipe Social. De acordo com o 3º Mapa de Negócios de Impacto, oito em cada 10 empreendedores utilizam tecnologias inovadoras em seus negócios, sendo 31% enquadradas em *Big Data*, ramo que engloba *Machine Learning* (PIPE SOCIAL, 2021). Os algoritmos de ML são altamente interdisciplinares, sendo capazes de classificar e achar padrões escondidos nos dados (OSISANWO, 2017), como será detalhado nos capítulos seguintes.

Já no campo de *DS*, técnicas de ML vêm sendo utilizadas em diversos campos, o que resulta em cada vez mais aplicações relacionadas aos campos social, ético e de privacidade. Com isso, as dúvidas quanto à imparcialidade dos códigos têm sido intensificadas, juntamente com as soluções para deixar as decisões mais claras e menos imparciais. O viés vai além de apenas uma falha no código, pois pode refletir discriminações sociais sérias, como racismo, preconceito relacionados à orientação sexual, xenofobia e exclusões de gênero ou de condição social, por exemplo. Sendo assim, a busca por soluções que mitiguem o viés algorítmico é de extrema importância para gerar soluções justas e éticas (ALVES; ANDRADE, 2021). Alguns exemplos práticos de vieses nos códigos de ML serão explicitados no capítulo 3.

Para entender como o viés algorítmico pode ser aplicado ao microcrédito, é importante ressaltar que hoje em dia existem diversas plataformas *online* que trabalham com a prática dos microempréstimos, como Kiva, Prosper e LendingClub. No caso da plataforma Kiva, por exemplo, o modelo de negócio é baseado no financiamento coletivo, no qual uma pessoa que deseja financiar uma iniciativa entra no site e consegue encontrar negócios de diferentes países e de diferentes ramos. O financiamento usualmente ocorre com um grupo de pessoas emprestando o dinheiro até que a meta do projeto seja atendida. No entanto, existe uma problemática acerca do alcance de determinados projetos apresentam, visto que o algoritmo do site se baseia em escolhas anteriores do usuário, fazendo com que muitas vezes os resultados das buscas sejam enviesados e excluam certos países, gêneros e tipos de negócios em específico. Em algoritmos utilizados para a promoção do microcrédito, o viés pode ocorrer ao favorecer empréstimos a usuários recorrentes e confiáveis de crédito, deixando em segundo plano os novos usuários, ou então quando os anúncios de novos empreendimentos para financiamento são exibidos a partir do histórico de doações do usuário, podendo exibir uma tendência baseada em pré-conceitos do mesmo (SARKAR;ALVARI, 2020).

1.3 ESCOPO DO TRABALHO

O estudo em questão apresenta como objeto de pesquisa modelos de ML que são utilizados no microcrédito. Foram avaliados modelos já classificados previamente por diversos autores de estudos e artigos entre os quatro tipos de aprendizado de máquina, conforme explicitado no capítulo 2. Foram alvo do estudo materiais já publicados devido à dificuldade de acesso aos códigos utilizados por empreendimentos sociais de microcrédito, além do fato de que alguns dos dados sensíveis e sofrem sigilo empresarial. Os códigos avaliados já se encontravam classificados, de modo que foram evidenciadas apenas suas características que se enquadram nas técnicas de aprendizado, não tendo sido necessária realização da classificação no presente trabalho.

Além disso, são exibidos modos de viés algorítmicos e suas consequências na sociedade, com o enfoque para a problemática enfrentada na área do microcrédito. A finalidade é realizar uma interseção entre os três temas de interesse: ML, empreendedorismo social e viés algorítmico.

Grande parte dos materiais analisados foram publicados entre 2017 e 2021, evidenciando a atualidade do assunto. No entanto, também há a presença de estudos em períodos anteriores, entre 2004 e 2016, o que pode indicar que o assunto perpetua entre os anos, e se renovando com as problemáticas do período no qual estão inseridos. Todavia, é importante destacar que o cerne dos estudos não se distingue devido à variação nas datas de publicação. Durante a pesquisa foi possível perceber a escassez de materiais que relacionem o microcrédito às técnicas de ML, além da falta de aplicações das ferramentas no microcrédito no Brasil. Ademais, não foram analisados os códigos de sites utilizados como exemplo, por serem dados sensíveis e não disponibilizados pelas instituições.

Sendo assim, a questão de pesquisa principal é: Como os mecanismos de ML aplicados em contextos de microcrédito podem ser afetados pelo viés algorítmico, e como as empresas podem lidar e solucionar esses problemas?

1.4 ELABORAÇÃO DOS OBJETIVOS

O trabalho tem como objetivo geral identificar os riscos e as potenciais soluções para situações de viés algorítmico em projetos de microcrédito que utilizam *Machine Learning*. Além disso, o trabalho apresenta os seguintes objetivos específicos:

- Definir os conceitos relacionados ao empreendedorismo social e *Machine Learning*, e fazer um recorte de como os mesmos podem reduzir as desigualdades sociais através do microcrédito;
- Identificar o papel do microcrédito dentro de iniciativas de empreendedorismo social;
- Classificar possíveis formas de viés algorítmico em processos de *Machine Learning*
- Descrever métodos de aprendizado adotado por sistemas de *Machine Learning*;
- Avaliar como cada método de aprendizado de ML pode ser usado para prevenir viés algorítmico.

1.5 DEFINIÇÃO DA METODOLOGIA

A metodologia do presente trabalho foi baseada na classificação de Gil (2010), analisando os métodos, instrumentos de pesquisa, abordagem e objetivos.

Por se tratar de uma análise de estudos já publicados e elaboração de uma hipótese da classificação dos mesmos, tornando o assunto mais explícito e familiar ao leitor, a pesquisa apresenta um caráter exploratório. Além disso, ao caracterizar os modelos de ML, empreendedorismo, viés algorítmico e explicitar o conceito de microcrédito, também pode-se dizer que a pesquisa apresenta caráter descritivo. Diante disso, de acordo com a classificação dos métodos de pesquisa de Gil (2010), a pesquisa em questão é uma combinação entre pesquisa exploratória e pesquisa descritiva.

Como explicitado no escopo do trabalho, serão avaliadas ferramentas de ML aplicadas ao microcrédito, além de buscar por possíveis soluções para o viés algorítmico, levando em consideração os estudos já existentes. Com isso, serão analisadas diversas publicações e materiais teóricos, o que permite um conhecimento amplo e detalhado de objetos, se enquadrando na modalidade bibliográfica (GIL, 2010).

Além disso, serão avaliadas as características do site Kiva.org, como exemplificação das informações analisadas, sendo considerado como estudo de caso (GIL, 2010).

Por fim, é um trabalho de natureza aplicada que, segundo Gil (2010), visa solucionar questões no âmbito social que o pesquisador está inserido, como é o caso do empreendedorismo social e o microcrédito na redução das desigualdades sociais.

1.6 ESTRUTURA DO TRABALHO

O presente trabalho é constituído por quatro capítulos, cuja estrutura é descrita abaixo.

O primeiro capítulo é composto pela introdução, contendo o tema definido, as justificativas para sua escolha, além da definição da problemática e a escolha da metodologia.

No capítulo 2 é exposta a revisão bibliográfica, que contextualiza os temas empreendedorismo social e *Data Science* aplicados a redução das desigualdades sociais, como microcrédito e *Machine Learning.* Para entender como os conteúdos tratados se interligam, é importante conhecer o que é o microcrédito, e como o mesmo consegue reduzir as desigualdades sociais. Para isso, no presente trabalho é feita uma breve conceituação, com forte enfoque na figura de Yunus, além da apresentação de exemplos de sucesso que reforçam a importância do artifício financeiro em questão.

O capítulo 3 é destinado à descrição da pesquisa e da metodologia adotada no trabalho. Já no capítulo 4 são expostas as conclusões acerca do assunto estudado e discutido.

2. EMPREENDEDORISMO SOCIAL, MICROCRÉDITO DATA SCIENCE E CONCEITOS CORRELATOS

2.1 EMPREENDEDORISMO SOCIAL

2.1.1 DEFINIÇÃO DE EMPREENDEDORISMO SOCIAL

Em um cenário no qual o Estado não consegue acompanhar as mudanças da sociedade e solucionar plenamente os problemas gerados pelas mesmas, principalmente no campo econômico, emergem iniciativas que partem da própria sociedade civil, dando margem para o empreendedorismo social (PARENTE, 2011).

As organizações sem fins lucrativos, que usualmente ficavam responsáveis por identificar as demandas sociais e resolvê-las, perceberam uma oportunidade de preencher as lacunas orçamentárias e passaram a transformar suas atividades em um negócio que trouxesse retorno financeiro. Cada parte do mundo abordou o empreendedorismo social de uma maneira diferente, percebendo os hiatos do mercado tradicional e criando um novo nicho. Na Ásia, pode-se citar, novamente, Muhammad Yunus e o Gramming Bank (MACHADO, 2020). Hoje em dia, ser um empreendedor social significa focar em solucionar problemas da comunidade, visando minorias que não são assistidas pelas ações governamentais ou não-governamentais, mas compactuando com o caráter inovador e dinâmico dos negócios empreendedores (PARENTE, 2011).

Pode-se definir teoricamente empreendedorismo como “análise de como, por quem e com que efeitos, as oportunidades para criar bens e serviços futuros são descobertas, avaliadas e exploradas” (SHANE e VENKATARAMAN, 2000, p. 218). Diferentemente do conceito citado na frase anterior, que é mais genérico para empreendimentos como um todo, o empreendedorismo social se baseia na inovação e na geração de valor (financeiro e humano) para o empreendedor e para a comunidade. Sua principal diferença com o empreendedorismo tradicional é a finalidade, ou seja, no objetivo principal (PARENTE, 2011).

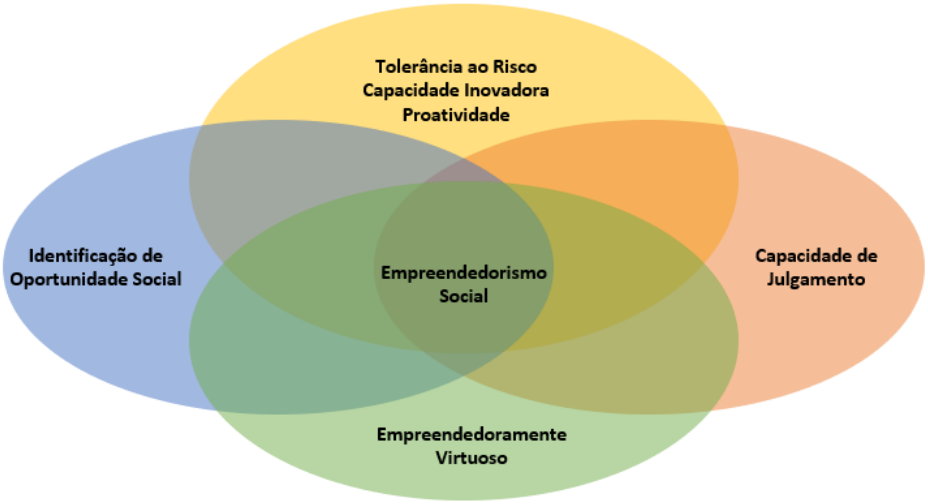
No caso do empreendedorismo tradicional, ele foca na acumulação de capital e lucro, enquanto no empreendedorismo social o foco é na geração de valor social, isto é, o objetivo é impactar e impulsionar a sociedade (PARENTE, 2011). Empreender socialmente pode ser entendido como levar as práticas empresariais para o campo da resolução de problemas sociais, utilizando de conceitos da inovação para buscar por soluções que o governo não consegue encontrar (MACHADO, 2020). Além disso, para se diferenciar de uma ONG, pode-se

considerar um empreendimento social uma organização na qual se obtém lucro, porém este é reinvestido em seus projetos sociais (CRUZ, 2012).

Existe uma discussão acerca do conceito de empreendedorismo social, visto que muitos teóricos entendem que o lucro deveria desqualificar uma organização de se enquadrar em tal ramo. No entanto, existe uma outra classe de pesquisadores, incluindo Muhammad Yunus, com a visão de que o lucro não exclui o lado social de um empreendimento, desde que seu objetivo primário seja o social (CRUZ, 2012).

Diante disso, pode-se definir o empreendedorismo social como um processo de diversas dimensões, que envolvem o lado empreendedor tradicional e o lado social, como a capacidade de julgamento e percepção das oportunidades de mercado, atreladas ao comprometimento com a sociedade, como pode ser identificado na Figura 2 (CRUZ, 2012):

Figura 2 - Teor multidisciplinar do empreendedorismo social



Fonte: Cruz (2012)

A esfera da “Identificação de Oportunidade Social” representa a capacidade do empreendedor de reconhecer e tirar proveito das situações na qual ele se encontra, descobrindo negócios potenciais. “Empreendedoramente Virtuoso” é o adjetivo que relaciona o empreendedorismo social à criação de valor para a sociedade na qual está inserido, demonstrando engajamento e dedicação a sua missão. Já os adjetivos “Tolerância ao Risco,

Capacidade Inovadora, Proatividade e Capacidade de Julgamento" se encontram em seu caráter literal (CRUZ, 2012).

Vale ainda ressaltar que, como já dito por Bornstein (2007), para se tornar um empreendedor social não é preciso ser alguém extraordinário, capaz de salvar o mundo. É preciso ter uma bagagem, capacidade de enxergar oportunidades usualmente invisíveis e, principalmente, engajamento com a sociedade na qual se está inserido.

Com as mudanças que o mundo dos negócios vem sofrendo, pode-se observar a tendência de surgimento de novos modelos de negócio que podem acabar transformando as organizações e a estrutura das comunidades. Nesse contexto, a importância da sustentabilidade em seu âmbito social, econômico e ambiental está cada vez mais em evidência (ÁVILA, 2016).

Diante disso, é possível citar a eclosão dos negócios de impacto social, que têm como principal característica a interseção entre o retorno econômico e o impacto social positivo (ÁVILA,2016), que acaba por ser sua principal distinção entre os negócios tradicionais (BRANDÃO; ÁRIDA;CRUZ, 2014).

Pode-se definir negócio de impacto social como uma organização economicamente rentável, que seja capaz de atingir positivamente a população de baixa renda, seja através da oferta de produtos e serviços a preços acessíveis, ou então através da inserção da mesma no mercado de trabalho, gerando emprego e renda. Assim como no empreendedorismo social, existe a característica de ser um modelo de negócio autossustentável, com a transformação social como foco principal (BRANDÃO; ÁRIDA;CRUZ, 2014). Para Scherer (2014), para se enquadrar em um negócio de impacto, uma organização deve introduzir a população de baixa renda na cadeia de valor, gerando empregos e renda.

Todavia, para outros autores como Yunus (2010), Kerlin (2009) e Naigeborin (2010) negócios de impacto social são uma categoria dentro do empreendedorismo social, considerando-os como iniciativas que são rentáveis economicamente e utilizam artifícios para melhorar a qualidade de vida da população (VIEIRA, 2019).

No presente trabalho, será utilizada a classificação mais ampla, na qual negócios de impacto social são uma categoria dentro do empreendedorismo social, além de dar enfoque em empreendimentos sociais voltados para o acesso a serviços financeiros, mais especificamente ao microcrédito.

2.1.2 EMPREENDEDORISMO SOCIAL NO MICROCRÉDITO

O empreendedorismo social vem surgindo como um fenômeno, como uma maneira de combater a pobreza, as desigualdades sociais e de reduzir os impactos negativos causados ao

meio ambiente. Nesse contexto, pode-se perceber que os empreendimentos sociais têm sido capazes de mudar e aprimorar socialmente as zonas nas quais estão inseridos (BOSE, 2012).

É possível definir o desenvolvimento social em três esferas: a esfera dos direitos civis e políticos, a dos direitos socioculturais e econômicos e a dos direitos relacionados ao meio ambiente e ao desenvolvimento coletivo (BOSE, 2012). Com isso, o principal objetivo dos empreendimentos sociais é promover a equidade social, reduzindo ao máximo as desigualdades e buscando potencializar os benefícios para a população mais carente (SACHS, 2004).

Um exemplo do empreendedorismo social atrelado ao microcrédito é o site Kiva.org. O site é baseado no financiamento coletivo através de micro empréstimos, com o objetivo de quebrar as fronteiras do financiamento tradicional e da distância geográfica. Através da plataforma, estudantes conseguem pagar por seus estudos, mulheres empreendedoras conseguem abrir seus negócios e fazendeiros conseguem investir em técnicas e maquinários para aprimorar seus negócios. O site, até dezembro de 2022, permitiu o financiamento de 1,82 bilhões de dólares, para 79 diferentes países, e a união de 2,1 milhões de credores a 4,5 milhões de pessoas que precisavam de financiamento (KIVA.ORG, 2022).

A instituição Kiva apresenta dois principais meios de obter receita. O primeiro é através de uma taxa opcional no empréstimo, na qual o credor escolhe doar a mais do que o necessário para o empréstimo. Se o usuário pretende doar $100,00, existe a opção de doar 10% a mais, por exemplo, resultando em um total de $110,00 finais. Isso ocorre pois o valor dos empréstimos é 100% convertido para os microempreendedores e para a compensação dos credores. A segunda forma de receita é referente aos juros acumulados nas contas bancárias. Como o Kiva recebe valores de todo o mundo e deposita os valores em contas bancárias, esses valores estão sujeitos a taxas de lucro, que são destinadas para pagamento das despesas essenciais ao funcionamento da empresa. Além disso, o Kiva cobra uma taxa trimestral para as instituições financeiras parceiras que administram os fundos da organização (FLANNERY, 2007).

A interface gerada pelo empreendedorismo social, os negócios de impacto social e as redes socioeconômicas podem e vêm contribuindo cada vez mais para o desenvolvimento e a gestão das comunidades. A autonomia que empreender agrega a causa social viabiliza decisões horizontais, trocas saudáveis e o sentimento de solidariedade, que tendem a potencializar o impacto e democratizar o enfrentamento de demandas sociais (DA SILVA; MOURA; JUNQUEIRA, 2015).

A restrição total ou parcial da população ao sistema financeiro é um dos muitos fatores que aprofundam a pobreza de uma sociedade, consequentemente, contribuindo para ampliar as

desigualdades sociais. O acesso aos serviços financeiros tradicionais, usualmente, apresenta custos elevados e pré-requisitos inalcançáveis para a parcela mais carente da população, se tornando algo que não atende as verdadeiras necessidades de quem realmente precisa do sistema financeiro. Com isso, fazem-se necessárias alternativas que fogem dos padrões pré-estabelecidos, e uma delas é o microcrédito (CERNEV; PROENÇA, 2016).

Nesse contexto, surge a figura de Muhammad Yunus, que se dedicou a minimizar a pobreza através do microcrédito, criando o Grameen Bank (ESTY,2011). O Grameen Bank foi pioneiro ao relacionar a concessão de crédito à concessão de empréstimos para pessoas pobres. Atendendo mais de 2,6 milhões de usuários, permite o acesso a melhores condições de vida, além de oportunidades de lutar por uma vida mais digna. O Sistema criado por Yunus, que hoje é amplamente utilizado ao redor do mundo, é flexível e age em diversas dimensões, impactando a vida das pessoas no âmbito social, político e econômico (LATIFEE, 2003).

Microcrédito é uma forma de concessão de empréstimos de pequeno valor, na casa dos US$ 100, tendo como público alvo pequenos empreendedores informais, microempresas, pessoas pobres, principalmente mulheres rurais de baixa renda, com pouco acesso ao sistema financeiro convencional. Usualmente, os valores são utilizados para geração de renda focada na economia informal (RAHMAN; ELAHI, 2006). Tal forma de concessão é importante pois democratiza o acesso à renda, movimentando uma parcela da economia que, sem esse apoio, seria inativa (CERNEV; PROENÇA,2016). O microcrédito se dá quando uma organização disponibiliza de maneira ágil a concessão de crédito para a população carente, com maneiras alternativas de garantia de pagamento, diferindo do modelo de financiamento tradicional. O principal foco é no mercado local, investindo em pequenos empreendedores, que são fidelizados a partir da liberação de mais crédito de acordo com a boa relação com a instituição. As taxas de juros cobradas são baixas, e os clientes são integrados através de agentes de crédito, gerando comprometimento mútuo entre ambos os envolvidos (RIBEIRO,2009).

Uma das principais diferenças entre o microcrédito e o serviço financeiro tradicional é que, ao invés de oferecer garantias como imóveis e títulos, para ser apto a receber o empréstimo, o cidadão precisa fazer poupanças regularmente (RAHMAN; ELAHI, 2006).

Na visão de Yunus, o crédito é um dos direitos básicos do ser humano, contribuindo para promover e alcançar condições de vida dignas. Permitir o acesso ao sistema financeiro pode acarretar na quebra do ciclo vicioso da pobreza, gerando renda, autoemprego, saúde, bem-estar e educação. É uma forma de diminuir a perpetuação das desigualdades sociais, acarretando no empoderamento de comunidades e, consequentemente, reduzindo vulnerabilidades intrínsecas (LATIFEE, 2003).

O microcrédito, no contexto social, emerge como uma possibilidade de transformar a realidade de milhares de pessoas em situação de pobreza ao redor do mundo. Ao oferecer micro empréstimos de maneira simples e menos burocrática para financiamento de negócios, dentre outros fins, estimula a geração de renda e empregos (SEGATTO; DIÓGENES; BISCAIA, 2012).

Em Bangladesh, no ano de 1982, Yunus iniciou sua jornada nos micro empréstimos com um valor de US $27,22 concedidos a 42 pessoas, mostrando que era possível emprestar pouco para muitos. Com isso, Yunus ganhou o Prêmio Nobel da Paz em 2006, e o Grameen Bank está presente em mais de 57 países (SEGATTO; DIÓGENES; BISCAIA, 2012). Ao emprestar a pequena quantia ao grupo de 42 mulheres, Yunus concluiu que pessoas carentes assumem e cumprem seus compromissos mesmo sem as garantias dos sistemas de crédito tradicionais, e que o acesso ao financiamento é imprescindível para uma sociedade se desenvolver sustentavelmente (PEREIRA; ROLO, 2012).

O acesso ao sistema financeiro representa algo que contribui para que o indivíduo esteja apto a se desenvolver e entender o seu potencial e o seu papel dentro de uma comunidade. Estudos recentes mostram fortes indícios de que o microcrédito auxilia no acesso aos bens de consumo, à construção de ativos, à educação e saúde de qualidade (LATIFEE, 2003).

No Brasil, o microcrédito vem crescendo através de ONGs, Cooperativas de Crédito, Organizações da Sociedade Civil de Interesse Público, Sociedades de Crédito ao Microempreendedor, programas de bancos oficiais, dentre outros (SEGATTO; DIÓGENES; BISCAIA, 2012). De maneira diferente aos demais países na qual as iniciativas do micro empréstimo partem da própria esfera civil, o microcrédito no Brasil tem partido de iniciativas públicas e/ou governamentais, estimulando o fornecimento de crédito à população de baixa renda (RIBEIRO,2009). De acordo com um estudo de caso realizado no Banco Comunitário Muiraquitã, a aplicação do microcrédito foi uma prática bem-sucedida, visto que a fração dos credores prósperos cresceu anualmente. Além disso, foi adotada uma política de troca de lixo pela moeda Muiraquitã e, no período do estudo, uma quantidade próxima de 10 toneladas de lixo foi reciclada. Ou seja, se aplicado de maneira adequada à sociedade, o microcrédito pode ir além da solução financeira, e pode afetar as soluções socioambientais (SEGATTO; DIÓGENES; BISCAIA, 2012).

É válido ressaltar que o microcrédito vai além de pequenos empréstimos. Com o microcrédito, a população pobre tem acesso a uma rede de cooperação entre microempresários, impulsionando o comércio local, gerando ajuda mútua para superação da pobreza e exclusão socioeconômica. É importante destacar, também, que o microcrédito pode se desenvolver como

um negócio social (cooperativas e sociedades de crédito) ou como outras formas de organização, como ONGs e bancos públicos (PEREIRA; ROLO, 2012).

Um exemplo de como o microcrédito pode contribuir para a comunidade é a moeda social Mumbuca, originária do município de Maricá. O investimento da prefeitura começou com moeda física, e se tornou uma moeda digital. Em seis meses de aplicação do projeto, já haviam 7.576 beneficiários, que recebiam uma quantia de R$70,00 que deveria ser reinvestida no comércio local, fortalecendo a dinamização da economia do município (CERNEV, 2019).

Além disso, no Brasil, temos experiências como o Banco da Mulher, com apoio da UNICEF, que é filiado ao *Women's World Banking* que, até meados de 2001, movimentou cerca de 9 mil créditos, valor equivalente a cerca de R$6,7 milhões. Pode-se citar também o VIVACRED, iniciativa do movimento Viva Rio, concedendo crédito a microempreendedores do Rio de Janeiro, totalizando mais de R$14 milhões até o final de 2001 (SIQUEIRA FILHO, 2005).

Pode-se evidenciar, também, a atuação governamental, que fomenta o microcrédito em escalas municipais e estaduais, juntamente com iniciativas privadas, concedendo crédito diretamente ao tomador. Tais iniciativas são conhecidas popularmente como "Bancos do Povo". Vale destacar a mais antiga rede de microcrédito do país, o sistema CEAPE, que atua em 13 estados brasileiros. O sistema é coordenado pelo CEAPE Nacional, que concede crédito e capacitação a microempreendedores locais (SIQUEIRA FILHO, 2005).

2.2 DATA SCIENCE

2.2.1 CONCEITO

Data Science, ou ciência de dados, é um termo que se tornou popular nos últimos anos. Mais do que a busca por uma definição exata do que DS faz, o foco atual é no que se consegue usufruir da mesma e qual é a sua constituição (ZHANG, 2021). A crescente geração de dados em tempo real e seu fluxo constante criou o desafio de descobrir o que pode ser feito com os mesmos, como aproveitá-los e aplicá-los utilizando seu máximo potencial. É na resolução desse desafio que entra DS (CURTY, 2016).

De acordo com Van Dyk (2015), existem três princípios básicos para *Data Science:*

1. gerência, agrupamento, organização e transformação de banco de dados;
2. utilização de ferramentas estatísticas e ML para transformar os dados em conhecimento e;

3. necessidade de infraestrutura computacional para realizar as análises de dados.

Pode-se dizer que *Data Science* é um modelo computacional para se descobrir padrões. Ou seja, é um sistema capaz de reconhecer regularidades, repetições e acumulação de elementos. Conhecer os padrões de mercado, por exemplo, é uma grande ferramenta de apoio à decisão, que vem sendo largamente utilizada e que cada vez ganha mais valor agregado (CAVIQUE, 2014).

Para Stanton (2012), *Data Science* envolve identificar, coletar, tratar, transformar e analisar grandes volumes de dados, que na grande maioria das vezes é heterogênea. Para isso, é preciso identificar as ferramentas mais adequadas para análise dos dados, além de como os resultados devem ser interpretados e exibidos para quem os interessa (CURTY; SERAFIM, 2016).

Além de saber atuar nas ferramentas e sistemas, um cientista de dados precisa ter habilidades e conhecimentos como modelagem de dados, pensamento analítico/lógico, linguagem de programação, banco de dados, plataformas *Open Source* (código aberto), visualização de dados, análise quantitativa/estatística, capacidade de monitoramento e controle e gerenciamento de projetos, dispositivos de armazenamento e conhecimentos gerais de *hardware*, etc. (CURTY; SERAFIM, 2016).

As ferramentas de *Data Mining*[2] (mineração de dados) muitas vezes são apresentadas como possíveis soluções de problemas sociais, aumentando a eficiência dos processos e sendo neutras às diferenças socioeconômicas. No entanto, ao utilizar os mecanismos de aprendizado de máquina, é necessário informar características e padrões nos quais os dados se enquadram, possibilitando a incorporação da subjetividade no código. Tal subjetividade é classificada como viés algorítmico, tendo maior ocorrência na definição de classes equivalentes, seleção de qualidades e categorias de modelagem, gerando uma interpretação da realidade baseada na visão do programador (SIMÕES-GOMES; ROBERTO; MENDONÇA,2020). Sendo assim, o programador precisa de habilidades não apenas voltadas para a computação, mas também para a percepção dos problemas que estão a sua volta, para evitar análises distorcidas dos dados (CURTY, SERAFIM, 2016). Assuntos e conceitos atrelados ao viés algorítmico serão aprofundados na seção 2.2.4.

---

[2] *Data Mining* é uma área dentro da DS que analisa os dados fornecidos e extrai-se padrões dos mesmos. Exemplos são a identificação de clientes com hábitos de compras iguais e encontrar características que levam a reclamações dos consumidores. Em *Data Mining* é de extrema importância a descoberta de padrões sequenciais (CAVIQUE, 2014).

2.2.2 CAMPOS DE APLICAÇÃO

Como *Data Science* pode ser considerado um conjunto de métodos utilizados na análise e no processamento de dados, auxiliando na tomada de decisões de maneira inteligente, seu campo de atuação é amplo e vasto (DOS PASSOS, 2016).

Dados são gerados desde pequenas ações, como a compra de um lanche em um restaurante *fast food* e utilização de redes sociais, como em ações mais sofisticadas, como investir na bolsa de valores. No entanto, não é só a geração de dados que aumentou significativamente, a utilização dos mesmos também vem crescendo de maneira exponencial. Para isso, a capacidade de processamento também precisou aumentar, juntamente com o armazenamento (CAVIQUE, 2014).

Atividades relacionadas a *Data Science* são utilizadas nas redes sociais como *Facebook* e *LinkedIn*, em soluções NoSQL[3]. No ramo de negócios, é possível identificar produtos mais vendidos, clientes com mais reclamações, identificar clientes com padrões de compras idênticos (clusterização), encontrar pontos comuns de reclamação etc. (CAVIQUE, 2014). No âmbito da Engenharia de Produção, podem ser utilizadas técnicas de melhorias de processos e diminuição de estoques através de *Data Mining* associado ao *Lean Manufacturing*[4] para identificação de padrões que definem a necessidade ou não de estoque, por exemplo (BRONSTRUP; SIMÕES, 2017).

Na medicina de precisão, pode-se citar o uso intensivo de DS, ou seja, na capacidade de conseguir segmentar a população e então prescrever medicamentos de acordo com a doença que o mesmo verdadeiramente funcione, por exemplo. Além disso, têm-se os prontuários eletrônicos, que permitem uma universalização dos dados do paciente, além da diminuição do espaço físico necessário para o armazenamento dos prontuários e receituários físicos (CHIAVEGATTO FILHO, 2015).

Levando em consideração o ramo da educação, *Data Science* pode auxiliar em mecanismos de previsão de reprovação, alerta de prazos, indicadores de desempenho, formação de perfis para recomendação de conteúdos, *feedbacks* personalizados, planejamento de atividades, compartilhamento de informações, etc. (TAKAKI; DUTRA, 2022).

---

[3] *NoSQL* é uma ferramenta que permite o armazenamento, tratamento e a consulta de dados de maneira extremamente eficiente, partindo de modelos racionais e de linguagem declarativa, sendo capaz de lidar com altos volumes de dados (CAVIQUE, 2014).

[4] *Lean Manufacturing* é uma metodologia baseada na eliminação de desperdícios para a diminuição dos custos da cadeia produtiva. O objetivo é efetuar mudanças que agreguem valor ao produto/serviço e eliminar as que não agregam. Para isso, é necessário criar valor, definir a cadeia produtiva, otimizar o fluxo de produção e aperfeiçoar o sistema, sempre monitorando as saídas e controlando o processo (PEREIRA, 2010).

Pode-se citar ainda a atuação de DS nas licitações governamentais, para escolher os vencedores de maneira imparcial e objetiva, levando em conta diretrizes lógicas e técnicas como regressão linear, máquina de vetores de suporte e regressão por vetores de suporte, por exemplo. Tais métodos são capazes de coletar, analisar e classificar grandes volumes de dados, identificando fraudes, desvios e auxiliando em processos de auditoria. (DE OLIVEIRA; DO RÊGO; DINIZ, 2019).

Os exemplos acima mostram que a DS cresceu em grande parte das áreas nos últimos anos, e que suas aplicações só tendem a se expandir com os avanços da tecnologia (CHIAVEGATTO FILHO, 2015).

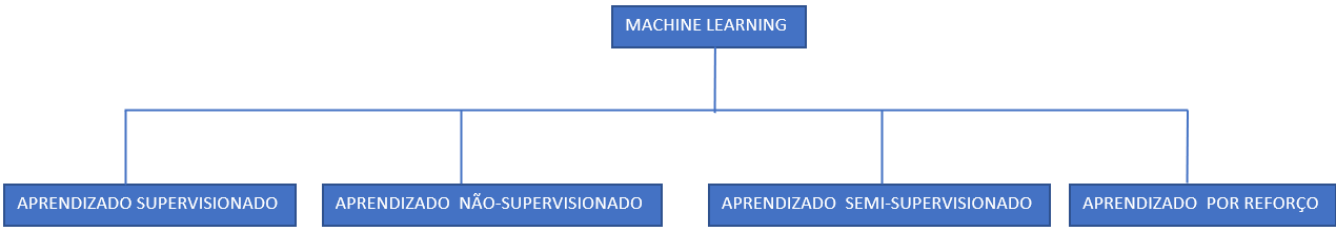
2.2.3 *MACHINE LEARNING*

Dentro da área de *Data Science*, existe um ramo que é focado na construção de algoritmos baseados no aprendizado de dados, e não em ações pré-programadas. Esse ramo é conhecido como *Machine Learning* (PAIXÃO, 2020). ML pode ser considerada uma das áreas com mais rápido crescimento e com mais vasto campo de atuação da ciência da computação. Os algoritmos gerados são altamente adaptáveis e com capacidade de aprendizado. (OSISANWO, 2017). Pode-se afirmar que ML é um método de repetição, que mapeia os dados e estabelece um padrão de rotulagem, de maneira confiável e precisa. (PAIXÃO, 2020)

Como grande parte dos modelos de DS é baseado nos “3Vs”: grande volume, alta velocidade e alta variedade de informações, a criação de métodos capazes de acompanhar esses dados é extremamente importante. Nesse contexto surge o ML, que é segmentado em três fases: pré-processamento, treinamento e avaliação (PAIXÃO, 2020). O que torna os algoritmos de ML sofisticados, é a capacidade de mapear e processar dados desbalanceados e heterogêneos (OSISANWO, 2017).

Na etapa de pré-processamento, é feita a organização do banco de dados, definindo o que deve ser analisado e separando-se os dados em treinamento e teste. No treinamento, o aprendizado pode ocorrer de maneiras diversas, que estão explicitadas na Figura 3 e serão mais detalhadas no decorrer deste tópico. Por último, na fase de avaliação do modelo, as diferentes maneiras de aprendizado são comparadas com os dados de teste e as saídas são geradas (PAIXÃO, 2020).

Figura 3 - Tipos de Aprendizado de *Machine Learning*



Fonte: Adaptado de Mahesh (2020)

Quando é feito aprendizado de maneira supervisionada, é necessária apresentação de exemplos de saída, ou seja, rótulos ou valores de destino para os dados que serão classificados (JANIESCH; ZSCHECH; HEINRICH, 2021). Os algoritmos aprendem a partir de conjuntos de exemplos rotulados, relacionando os atributos de entrada e de saída (PAIXÃO, 2020). Para exemplificar o modelo de aprendizado supervisionado, é a classificação de pessoas entre veganas e não veganas, cujos atributos são pré-estabelecidos e apresentam características específicas, como comer ou não carne, que conseguem ser identificados pelos algoritmos (MANESH, 2020).

Já no aprendizado não-supervisionado, não são informados rótulos pré-estabelecidos para o algoritmo, cuja função é encontrar os padrões em grupos de dados com características semelhantes, visando encontrar atributos ainda não conhecidos (revisão e aplicabilidade). Diferentemente do aprendizado supervisionado, não há respostas corretas ou erros, o programa aprende com os dados (MANESH, 2020).

O método de aprendizado supervisionado sofre maior interferência humana, se comparado com o não-supervisionado, porém demonstra maior acurácia e mais rapidez. O aprendizado não-supervisionado apresenta custo elevado e requer grandes quantidades de dados, porém é um método eficaz na redução da dimensionalidade do problema. Exemplos de algoritmos supervisionados são árvores de decisão e redes neurais, enquanto de não-supervisionados são agrupamento hierárquico e análise linear de discriminantes (PAIXÃO, 2020).

O método de treinamento semi-supervisionado é uma interseção entre os métodos supervisionado e não supervisionado, apresentando algumas características de ambos. O modelo utiliza tanto dados rotulados previamente quanto dados não rotulados, porém o objetivo é classificar os mesmos da melhor maneira com o mínimo de rótulos de treinamento possíveis (NEGRETTO, 2016).

Diferentemente dos modelos de supervisão, o aprendizado por reforço é baseado na tentativa e erro, no qual é feito um panorama geral do estado dos dados, as ações são limitadas e uma meta é estabelecida, para que então o algoritmo experimente e chegue às conclusões por conta própria (JANIESCH; ZSCHECH; HEINRICH, 2021). Há um agente, no caso o algoritmo, que interage diretamente com os dados por meio de ações e percebe os efeitos dessas ações por meio de recompensas (tentativa e erro), e o objetivo do agente é maximizar as recompensas, ou seja, os acertos (LEITE, 2012). O modelo de aprendizado por reforço é muito utilizado em jogos e no mercado eletrônico (JANIESCH; ZSCHECH; HEINRICH, 2021). O modelo não precisa de treinamento prévio e é altamente adaptável, sendo uma ferramenta útil em técnicas de codificação de ondas de rádio, por exemplo (LEITE, 2012).

2.2.4 MODELOS DE ANÁLISE DAS SAÍDAS

Conforme explicitado no tópico anterior, existem quatro métodos para aprendizado de máquina: supervisionado, não supervisionado, semi supervisionado e por reforço. Dentro dessas possibilidades, são obtidas diferentes saídas de dados, que devem ser analisadas de acordo com o objetivo do programador e do usuário.

Diante disso, surgem diferentes modelos para conseguir extrair as características necessárias dos dados de saída. No presente trabalho serão abordados os métodos de classificação, regressão, associação e agrupamento.

Nos problemas de classificação, usualmente são utilizados algoritmos de aprendizado supervisionado, nos quais são inseridos no código suficientes exemplos de entrada e saída, para que o mesmo categorize as informações com base nestas informações. O programa tem como objetivo aprender uma regra que entregue como saída um mapeamento correto das entradas dentro de classes pré definidas. Nesse tipo de problema, o objetivo é sempre categorizar. Um exemplo é na liberação de crediário para um cliente de varejo. Na base de dados, existem diversas características de clientes prévios, que definem se o mesmo será um bom ou mau pagador. O objetivo do programa é definir se um novo cliente será um bom pagador e pode receber o crédito. A Figura 4 abaixo representa o problema. De acordo com as características da base de dados, o código definirá quais são os atributos de um bom pagador e classificará “João” como sendo um, ou não (ESCOVEDO; KOSHIYAMA, 2020).

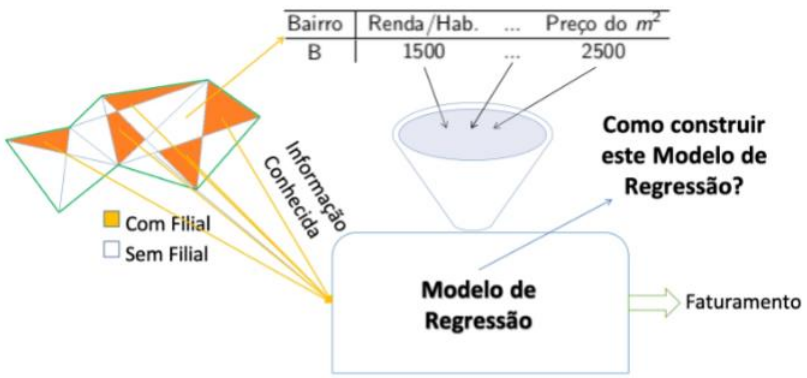
Figura 4 - Problema de Liberação de Crédito

Cliente	Renda	Idade	Est. Civil	Pagador
Ana	360	25	C	MP
Bernardo	350	33	S	MP
Carlos	1100	56	S	MP
...	...	...	...	...
Xavier	600	33	S	MP
Zé	800	29	S	BP
João	**750**	**38**	**C**	**?**

Fonte: Escovedo e Koshiyama (2020)

Ainda de acordo com Escovedo e Koshiyama (2020), uma segunda categoria de problemas é a regressão. Neste caso, diferente da classificação, que retorna saídas categóricas, a regressão apresenta saídas numéricas. De maneira similar, na regressão o método de aprendizado usual é o supervisionado. Problemas de regressão avaliam dados históricos e retornam quais dados de entradas vão apresentar características similares. Um exemplo é uma empresa que quer abrir uma nova filial, porém não sabe qual o melhor local para abertura. As informações de faturamento de locais que já apresentam filiais serão utilizadas como base, e servirão para estimar qual seria o faturamento em uma localidade desconhecida. As Figuras 5 e 6 abaixo exemplificam como o modelo foi pensado e alguns dos dados avaliados, respectivamente.

Figura 5 – Modelo de Regressão



Fonte: Escovedo e Koshiyama (2020)

Figura 6 – Dados do Modelo de Regressão

	Bairro	Renda/Hab.	...	Preço do m^2	Faturamento
Observado	A	1500	...	2500	105.000
	H	2400	...	5300	180.000
	...	...	...	...	...
	L	3400	...	2750	150.000
Esperado	B	2500	...	1780	125.000
	C	**1000**	...	**3500**	**200.000**
	D	4300	...	6500	
	...	...	...	...	...
	K	7000	...	8900	180.000
	M	2800	...	3900	120.000

Fonte: Escovedo e Koshiyama (2020)

Por fim, serão apresentados dois tipos de problemas que apresentam como base o método de aprendizado não supervisionado (entradas não apresentam rótulos de saída associados). O objetivo de ambos é reconhecer características importantes e em comum para os dados.

Nos problemas de associação, a finalidade é avaliar a primeira escolha do usuário e prever a segunda. No caso de vendas de produtos online, por exemplo, o alvo é identificar quais devem ser as sugestões para os clientes com base em suas escolhas anteriores. Busca-se assegurar o quanto uma ocorrência antecedente pode afetar a ocorrência consequente.

Por último, temos os problemas de agrupamento, que categorizam os dados dentro de subconjuntos (grupos). São problemas que visam agrupar elementos com propriedades em comum, que os aproximem entre si e os diferenciem de outros grupos. Os objetos de um grupo precisam ser mais similares entre si do que com os objetos de outros grupos.

3. VIÉS ALGORÍTMICO E O EMPREENDEDORISMO SOCIAL COM ENFOQUE NO MICROCRÉDITO

3.1 VIÉS ALGORÍTMICO E SEUS RISCOS

Com o aprimoramento das técnicas de *Data Science (DS) e Machine Learning (ML)*, a automatização das atividades têm otimizado e facilitado diversas áreas cotidianas das sociedades. Alguns exemplos a serem citados são as recomendações de buscas, atendimentos

automáticos, diagnósticos médicos e o reconhecimento facial em dispositivos inteligentes (ALVES; ANDRADE, 2021).

No entanto, os mecanismos de DS dependem dos modelos que são baseados, muitas vezes sendo simplificações de atividades diárias. Os programadores precisam definir quais são os dados de entrada, quais devem ser as saídas, além de, muitas vezes, definir como os dados serão processados. Isso pode fazer com que o código reflita as prioridades do programador e suas concepções, gerando pontos cegos aos algoritmos e transferindo para eles a subjetividade de seu criador (NUNES, 2018).

Sendo assim, o viés algorítmico pode ser traduzido como uma preferência humana por certos valores, devido ao seu contexto sociocultural. A preferência não se dá somente pelos dados de entrada e saída, mas principalmente pelo treinamento enviesado das técnicas de ML, fazendo com que as respostas deixem de ser neutras. O viés pode vir a privilegiar grupos de pessoas, intensificar preconceitos (de raça, gênero, orientação sexual etc.). Como exemplos, podem ser citados os casos de um programador negro que relatou ter suas fotos classificadas como de gorilas pelo Google Fotos; uma investigação do Estado de Nova York que identificou a concessão de limites diferenciados para homens e mulheres pelo cartão de crédito da Apple e a ferramenta norte americana COMPAS (*correctional offender management profiling for alternative sanctions*, em português, perfil de gerenciamento corretivo de infratores para sanções alternativas), no qual o código tinha a tendência de classificar negros como “prováveis reincidentes” e brancos como “baixo risco de reincidência” (ALVES; ANDRADE, 2021).

Em algoritmos utilizados para a promoção do microcrédito, o viés pode ocorrer ao favorecer empréstimos a usuários recorrentes e confiáveis de crédito, deixando em segundo plano os novos usuários. Além disso, em plataformas que permitem a promoção de anúncios de empreendimentos sociais de pessoas carentes a grupos que desejam financiar algum dos mesmos, existe o viés algorítmico quando os sites exibem anúncios de acordo com a visão de quem financia. Ou seja, os códigos analisam os perfis para os quais o usuário já emprestou alguma quantia, e exibem novos perfis de acordo com características similares entre ambos. Isso pode fazer com que os novos perfis exibidos favoreçam empreendimentos de um determinado grupo, país, ou de um gênero específico, tudo isso baseado nos pré-conceitos do usuário (SARKAR; ALVARI, 2020).

Os algoritmos de ML, tradicionalmente, eram elaborados e alimentados por cientistas de dados, contendo informações mais confiáveis e previsões mais assertivas. Hoje em dia, com o avanço da tecnologia e o acesso à informação, os programas vêm recebendo dados da população em geral, o que pode tornar as saídas tendenciosas e que fogem da realidade. Nesse

contexto, surge o termo Viés Algorítmico (SUN; NASRAOUI; SHAFTO, 2020). A segregação ocorre quando os *inputs* são contaminados e produzem distorções nos *outputs* (DA ROCHA; PORTO; ABAURRE, 2020), ou seja, quando o programador utiliza mecanismos para suavizar tendências, gerando respostas manipuladas, ou quando a própria entrada dos dados favorece tal tendenciamento.

No entanto, é importante entender o conceito do viés, pois sua interpretação varia dependendo de sua ocorrência nos algoritmos, assim como suas possíveis soluções. O viés pode ser observado de diversas maneiras no universo da DS. Um tipo de viés a ser citado é o de treinamento (do algoritmo), no qual as informações de entrada ou de treinamento fornecidas representam uma tendência do programador ou do usuário, transmitindo seus preconceitos e preferências ao modelo. Essa forma de viés muitas vezes passa despercebida, visto que programadores raramente divulgam a base de dados completa utilizada para o treinamento, além do fato de que o usuário muitas vezes não percebe o viés, já que as informações de saída estão de acordo com seu posicionamento e visão. Um segundo método de viés é o de processamento. Como a tendência surge do próprio código, também pode ocorrer de o programador utilizar seu conhecimento e as ferramentas de ML para mascarar resultados e alterar as saídas, manipulando os dados que serão repassados para o público. Uma terceira forma de viés é o de foco algorítmico, em que são utilizadas categorias de classificação irrelevantes, deixando de lado informações estatisticamente valiosas, alterando as saídas do programa. Pode ocorrer também de existirem categorias ilegais, que não podem ser utilizadas para análise. Os três tipos de vieses supracitados dizem respeito a questões computacionais. Os próximos dois casos já são fundamentados na utilização incorreta dos algoritmos (DANKS; LONDON, 2017).

O viés de contexto de transferência é baseado nas inconsistências entre o código e o contexto no qual o mesmo é aplicado. Para gerar um modelo de classificação, é necessário inserir informações iniciais para que o computador entenda o que precisa ser feito. Esses parâmetros iniciais seguem um contexto, que trará resultados corretos quando os dados avaliados estão inseridos nesse mesmo contexto. Caso contrário, as saídas serão uma distorção da realidade. Para exemplificar, pode-se citar um algoritmo para monitoramento de saúde de uma clínica urbana sendo utilizado em uma clínica rural. Os conceitos estatísticos utilizados na cidade provavelmente não se aplicarão para o padrão de vida de uma comunidade rural, resultando em dados não confiáveis. Por fim, surge o viés de interpretação, que depende do usuário, e não do código. Tal situação ocorre quando o usuário apresenta conceitos e valores que diferem do código e de suas saídas. Nesse caso, a informação trazida pelo algoritmo será

interpretada de maneira enviesada, podendo ser distorcida por quem a analisa (DANKS; LONDON, 2017).

Para que o conceito de viés algorítmico fique mais claro e contextualizado, seguem alguns exemplos de situações reais nas quais ficou evidente a tendência gerada pelo código. Os exemplos mostram que o algoritmo é capaz de reproduzir os preconceitos de raça e de gênero já existentes na sociedade (ALVES, 2021). Com o aumento da escalabilidade dos algoritmos, a maneira como a realidade é vista pela população e suas formas de interação são afetadas pelos preconceitos dos códigos e programadores. Nas redes sociais, o efeito é o mesmo. Os conteúdos são entregues para a população baseados em algoritmos que os classificam como relevantes ou não para o público. Tal classificação se dá pela observação do que é consumido ou negado pelo indivíduo, entregando conteúdos similares ao que é aceito, e não entregando informações parecidas com o que foi negado. De maneira superficial, muitos entendem que esse é um mecanismo justo e que melhora a qualidade da experiência do usuário. No entanto, pode-se observar que as etapas seguidas pelos códigos, por mais que sejam matemáticas e estatísticas, não estão isentas do preconceito racial (DA ROCHA; PORTO; ABAURRE, 2020).

De acordo com Da Rocha, Porto e Abaurre (2020, v.1, p. 13):

> "[...] O que gera, de fato, a produção de resultados não-inclusivos é o problema de dados de treinamento e programação de rotinas autoexecutáveis enviesados: algoritmos treinados em um conjunto de dados do mundo real que refletem a discriminação existente podem muito bem replicar essa discriminação, sobretudo na indústria de criação de conteúdo digital que indissociavelmente se deslinda nas plataformas das redes sociais virtuais."

Se essa questão não for levada em consideração, as chances de banalizar e enraizar ainda mais o racismo na sociedade e em suas atividades cotidianas é muito grande. Os consumidores das redes sociais são capazes de criar, modificar e compartilhar tendências, alterando os dados que antes eram manipulados apenas por programadores. Com isso, ao se analisar os números de pessoas atingidas e remuneração, existe uma disparidade entre criadores de conteúdos brancos e pretos. Isso não ocorre somente no retorno das plataformas, mas também no pagamento que grandes empresas oferecem para a realização de publicidades nas redes sociais (DA ROCHA; PORTO; ABAURRE, 2020). Em uma pesquisa realizada em Oxford (SAKO; PARNHAM, 2021), foi possível verificar que mulheres empreendedoras recebem menos investimentos de acionistas privados, simplesmente pelo seu gênero. Além de enfrentar a resistência masculina em investir em empresas lideradas por mulheres, existe a possibilidade da inexistência de neutralidade nos algoritmos que classificam empreendimentos quanto à concessão de crédito. Um empresário no campo tecnológico afirmou que recebeu um

limite de crédito equivalente a 20 vezes o limite concedido para sua esposa, mesmo que a renda dela fosse maior que a dele. Tal situação mostra que os métodos utilizados em *scores* bancários podem originar discriminações e diferenças nas oportunidades oferecidas para homens e mulheres (SAKO; PARNHAM, 2021 apud ALVES; ANDRADE, 2021).

Por fim, fica visível que o viés algorítmico pode afetar a confiabilidade dos dados, gerando cenários que são distorções da realidade. É preciso ser cauteloso e atento para que os códigos não reflitam as preferências do programador e dos usuários, visto que a máquina deve ser uma ferramenta de apoio à decisão mais assertiva e menos subjetiva.

3.2 *MACHINE LEARNING*: DIFERENTES FORMAS DE APRENDIZADO E O VIÉS ALGORÍTMICO

Para entender como o viés algorítmico pode afetar os diferentes códigos de ML, é importante que fiquem claras algumas possibilidades de aplicação dos tipos de algoritmo supracitados.

3.2.1 APRENDIZADO SUPERVISIONADO

O primeiro método de treinamento abordado será o supervisionado, que consiste na avaliação e categorização de dados através de um conjunto de variáveis pré-definidas. Tais variáveis, que podem ser consideradas como as entradas do programa, apresentam influências sobre as saídas. Sendo assim, o objetivo é prever as saídas e acordo com os dados de entrada. As saídas são os dados que dependem das variáveis preditoras. É uma forma de generalização capaz de gerar prognósticos (NABINGER, 2018).

Um exemplo de aplicação do aprendizado supervisionado em assuntos cotidianos é o estudo de Nabinger (2018), que utilizou o algoritmo para caracterizar os resultados da Copa do Mundo de Futebol de 2018. Para isso, o autor adotou diversos modelos baseados no método supervisionado, que visavam prever os resultados dos jogos. Os dados utilizados no estudo foram retirados do site da FIFA, levando em consideração informações sobre quantidade de gols, finalizações, posse de bola, passes, bloqueios, cartões amarelos, momento do gol, setor do campo, pênaltis, tipo de clima e prorrogação. Com esses dados, foi elaborado um modelo de regressão com variáveis de entrada que foram classificadas através de valores atribuídos às mesmas. As saídas dos modelos foram comparadas aos resultados reais dos jogos. Como saídas, o autor indicou o modelo que melhor previu os resultados dos jogos, que pode ser usado para prever partidas de diversos campeonatos. Dos nove modelos testados no estudo, sete

apresentaram taxas de acerto acima de 50%, e dois ficaram abaixo da média. O modelo com melhor desempenho apresentou uma taxa de acerto de 73,44%, enquanto o de pior desempenho obteve uma taxa de acerto de 37,5%.

Outro exemplo para a aplicação do método em questão é a mitigação de ataques em rede a servidores, conforme observado no estudo de Silva e Schaeffer-Filho (2021). Os autores buscaram otimizar os custos operacionais ao contratar uma empresa que fornecesse a identificação de fluxos maliciosos através de um modelo supervisionado. Para isso, o código precisava ser capaz de perceber ataques externos e selecionar uma aplicação que fosse capaz de mitigá-los. O programa recebeu como variáveis as características de ações benignas no sistema (manutenções, melhorias e correções) ou ataques malignos, que as combinam de maneira a detectar anomalias e mitigações. Caso os dados recebidos se classificassem dentro das características de um ataque maligno, era proposta uma solução para o problema, reduzindo os gastos da empresa com ações corretivas.

Com isso, é possível perceber que a aplicação do modelo supervisionado pode se dar em diversos campos, indo de temáticas simples como a previsão de partidas de futebol até a segurança cibernética de uma companhia, demonstrando seu teor disciplinar e sua relevância para o cotidiano da população.

No entanto, vale ressaltar que o modelo supervisionado é o que mais sofre a interferência humana, pois os dados são classificados de acordo com categorias inseridas pelo próprio programador ou usuário, podendo resultar em saídas generalizadas por pré-conceitos já existentes. Com isso, a imparcialidade da máquina pode ser fortemente afetada pelos hábitos do programador/usuário.

Sendo assim, é importante que sejam estudados métodos para reduzir ou até mesmo eliminar todos os efeitos do viés, pois estes podem trazer grandes problemas e custos para os usuários. No caso de Nabinger (2018), resultados errados não apresentam consequências severas para a população. Já no estudo de Silva e Schaeffer-Filho (2021), caso as classificações sejam feitas de maneira errada e mascarando a realidade, as empresas podem sofrer ataques cibernéticos criminosos, resultando no vazamento de dados confidenciais e informações sensíveis, o que acarreta perda da confiabilidade e também gastos para consertar os erros e possíveis problemas judiciais.

3.2.2 APRENDIZADO NÃO SUPERVISIONADO

O segundo método a ser abordado é o não supervisionado. Nesse caso, diferentemente do treinamento supervisionado, no qual as saídas são baseadas nas entradas preditivas e em uma resposta de interesse, é usualmente utilizado para gerar regras de predição e compreender a configuração dos dados. Ou seja, em vez de classificar os dados de acordo com variáveis informadas pelo programador, as saídas serão estimativas de ocorrência de características comuns descobertas pelo programa (IZBICKI et al., 2019).

Tal método também pode ser amplamente utilizado no ramo da medicina, com o intuito de prever doenças e suas complicações. Um exemplo é o estudo de Pires et al. (2021), que utilizou técnicas de ML para mapear o comportamento do cérebro em pacientes com epilepsia. Os autores utilizaram uma base de dados disponibilizada pelo MIT que consiste em sinais enviados pelo cérebro nos pacientes da pediatria. Os resultados mostraram como é o comportamento do cérebro durante as crises epiléticas, mostrando a probabilidade das mesmas durante diferentes períodos da doença.

Outro exemplo, ainda no campo da medicina, é o material publicado por Izbicki et al. (2019). Nele, existem vetores de mensuração, para realizar as predições, e vetores para as respostas de interesse, que no caso são baseados no comportamento dos indivíduos pesquisados, que devem fornecer a expectativa de vida quando idosos. O modelo definiu valores para os parâmetros observados, considerando a possibilidade de falsos positivos e negativos de previsão de óbitos, conforme a Figura 7.

Figura 7 – Matriz de Valores para a Resposta Predita

		Resposta observada		
		Presente	**Ausente**	
Resposta predita	Presente	Verdadeiro positivo (*a*)	Falso positivo (*b*)	*a + b*
	Ausente	Falso negativo (*c*)	Verdadeiro negativo (*d*)	*c + d*
		a + c	b + d	

Fonte: IZBICKI et al., 2019

Os dados analisados foram obtidos nas saídas do estudo SABE, com indivíduos com mais de 60 anos e residentes em São Paulo. Os pesquisadores foram até o domicílio dos pesquisados e seguiram um roteiro de perguntas. Os idosos foram entrevistados inicialmente

em 2000, sendo abordados novamente em 2006 e 2010. Com isso, os pesquisadores conseguiram acompanhar os entrevistados e entender como a saúde dos mesmos mudou ao longo dos 10 anos, englobando também os óbitos no período. Para a execução do código de ML, foram selecionadas 37 variáveis, incluindo idade, sexo, escolaridade, hábitos de tabagismo, consumo de bebidas alcoólicas, rotina de alimentação, comorbidades, condições de mobilidade e estado funcional. Com isso, foi possível encontrar as variáveis que mais influenciam na expectativa de vida dos idosos em questão (IZBICKI et al., 2019).

À vista disso, fica cada vez mais evidente que as técnicas de ML são multidisciplinares e atuam em áreas nas quais o erro e o enviesamento podem afetar a qualidade de vida das pessoas, como a medicina.

Sobre o viés intrínseco ao treinamento não-supervisionado, as chances de replicar tendências sociais são menores se comparadas ao treinamento supervisionado. Como os dados não são rotulados através de categorias pré-definidas, e sim através da ocorrência dos eventos, as chances de o resultado ser mascarado pelas percepções e hábitos do programador/usuário são menores.

No entanto, ainda há a possibilidade do viés, visto que a base de dados usualmente é inserida e tratada pelo programador/usuário, que pode deixar de lado algumas informações cruciais para a saída correta. Na medicina, caso exemplificado acima, pode ocorrer o viés de treinamento, no qual o programador categoriza as informações de acordo com suas percepções e opiniões, e não com fatos comprovados pela medicina (como acreditar que o hábito de fumar não prejudica os pulmões, mesmo que já tenha sido provado). Podem, também, ser considerados os vieses de processamento e foco algorítmico, nos quais os programadores podem deixar de utilizar informações importantes, como um hábito do idoso, ou então criar categorias que são irrelevantes para a saúde do entrevistado, respectivamente, além do emprego de modelos matemáticos e estatísticos não condizentes com as questões exploradas. Por fim, os vieses de interpretação de dados também podem surgir, quando as saídas são utilizadas em contextos que diferem da amostra, ou então com erros de interpretação.

No caso de Izbicki et al. (2019), por exemplo, podem existir informações sensíveis sobre os pesquisados que não foram consideradas durante a programação e execução do modelo. Tais informações, caso sejam capazes de afetar a qualidade de vida do indivíduo, sendo deixadas de fora podem gerar saídas incertas, afetando a estimativa de expectativa de vida do mesmo. Já no caso de Pires et al. (2021), caso alguma saída seja avaliada por um médico que não apresenta conhecimentos aprofundados nos ataques epiléticos, podem transcorrer distorções no real estado de saúde do paciente.

3.2.3 APRENDIZADO SEMI SUPERVISIONADO

O penúltimo modelo de aprendizado é o semi supervisionado. Nesse caso, existe uma interseção entre os modelos supervisionado e não supervisionado. Uma parte dos dados é rotulada, conforme o aprendizado supervisionado, e outra parte é utilizada juntamente com os dados não rotulados para realizar o treinamento da máquina.

O objetivo é gerar saídas mais precisas se comparadas aos outros métodos, visto que não se restringe apenas às rotulações inseridas pelo programador (DAMBROS, 2019). Tais modelos usualmente são variações dos dois modelos citados nos tópicos anteriores, e suas aplicações são semelhantes aos mesmos (MATSUBARA, 2004).

O estudo de Matsubara (2004), por exemplo, gera uma classificação de textos de uma biblioteca, resultando em diretórios específicos de cada assunto identificado pelo algoritmo. Já o estudo de Dambros (2019) é, novamente, um exemplo de ML na medicina, no qual é utilizado o aprendizado de máquina para avaliar se componentes do DNA são funcionais ou não, prevendo se os mesmos afetarão ou não a tradução do miRNA[5] mensageiro.

Já no caso do aprendizado semi supervisionado, as chances do viés são uma combinação dos dois métodos supracitados. Como parte dos dados é treinada de maneira supervisionada, pode ocorrer o viés conforme explicitado no tópico 3.2.1, e a outra parte com viés semelhante ao exposto no tópico 3.2.2. Diante disso, o quanto o código pode ser comprometido depende do quanto de cada método será utilizado, e do quanto cada um deles já está afetado pelos possíveis vieses.

3.2.4 APRENDIZADO POR REFORÇO

Por fim, é exibido o cenário de aprendizado por reforço, que apresenta características distintas dos métodos supracitados. Em tal caso, as técnicas utilizadas são baseadas na tentativa e erro, definindo recompensas para o sucesso na tomada de decisão e penalidades para resultados sem sucesso (OTTONI; NEPOMUCENO; OLIVEIRA, 2016). O programa, através de uma função objetivo, aplica técnicas aos dados e vai aprendendo a melhor maneira de se prosseguir através dos sucessos e fracassos obtidos (JÚNIOR; WILTGEN; SANTOS, 2010).

---

[5] miRNA é uma expressão utilizada para se referir aos microRNAs, que são pequenas moléculas capazes de inibir a expressão genética. Ou seja, são estruturas celulares que inativam características genéticas, podendo evitar a intercorrência de doenças e falhas genéticas (JORGE; PEREIRA; OLIVEIRA; FERREIRA; MENON; DINIZ; PEZUK, 2021)

Tal método é amplamente utilizado em diversas áreas, como pesquisa operacional, controle de robôs, logística e controle semafórico.

A título de conhecimento, pode ser citado o estudo de Ottoni, Nepomuceno e Oliveira (2016), que comparam o melhor algoritmo para a resolução do problema do caixeiro viajante. O problema em questão visa encontrar a menor rota entre um conjunto de cidades. Para isso, o caixeiro é considerado o agente que precisa otimizar sua rota. São definidos os parâmetros como a distância e a média das distâncias entre as cidades, para que o algoritmo encontre o melhor caminho. Para que o programa encontre a melhor rota, são definidas as ações que o caixeiro pode executar, que consistem na intenção de ir de uma localidade para outra. Evitando que as rotas sejam repetidas, todas as ações que levem a locais já visitados estarão indisponíveis. Na avaliação da quilometragem entre um local e outro, o programa calcula as distâncias e subtrai de 1. Sendo assim, quanto mais negativa a saída, maior é a distância. Diante disso, o objetivo é encontrar o menor caminho, para que a penalidade seja cada vez menor.

Outro exemplo a ser citado é o de Bazzan (2021), que utiliza o aprendizado por reforço para estabelecer o controle semafórico de uma cidade. Bazzan (2021) considera os semáforos como agentes que aprendem, e que precisam encontrar a melhor rota entre os pares de origem-destino. Com isso, o programa entende que, quanto melhor for a integração entre os semáforos, mais rápido será o caminho. O objetivo é evitar as punições (perda de tempo com semáforos fechados) e encontrar a melhor rota entre os locais. Diferentemente do estudo de Ottoni, Nepomuceno e Oliveira (2016), o agente não é dotado de modelos que estipulam o comportamento do ambiente. Ou seja, o agente aprende a lidar com os cenários através das observações e dos direcionamentos informados para cada situação. Com isso, foi possível evidenciar os ganhos de tempo em viagens e a eficiência do sistema semafórico como um todo.

No caso do aprendizado por reforço, no qual a máquina aprende através da tentativa e erro, as chances de interferência são ainda menores que os outros três métodos supracitados. As chances de viés se dão através da inserção da função objetivo, na qual o programador/usuário pode introduzir modelos que excluem alguns parâmetros que podem ser importantes para a decisão da máquina.

No estudo de Ottoni, Nepomuceno e Oliveira (2016), como o programa recebe um modelo que estipula alguns dos comportamentos do ambiente, e tal modelo é feito pelo programador/usuário, as chances de respostas enviesadas são maiores do que no estudo de Bazzan (2021), no qual o programa não recebe nenhum molde, aprende realmente com seus erros e acertos (que são baseados no tempo de locomoção entre a origem e o destino).

Desse modo, é finalizada a exemplificação de utilização de cada um dos modelos definidos e suas chances de viés, para que no próximo tópico seja avaliada a aplicação dos códigos no setor de microcrédito.

3.3 POSSIBILIDADES E CONSEQUÊNCIAS DO VIÉS ALGORÍTMICO NO MICROCRÉDITO

Os modelos de aprendizado e seus possíveis vieses também são aplicáveis aos empreendimentos voltados para o microcrédito. Para entender o comportamento das populações e ofertar empréstimos de acordo com suas necessidades, usualmente são feitas pesquisas com voluntários, o que pode ser perigoso para a análise das estatísticas.

Um exemplo que pode ser relatado é na amostra que responde os questionários propostos. É pouco provável que a amostra seja um recorte exato de como uma comunidade se comporta, não tendo representantes de todas as categorias e classes sociais. Com isso, os algoritmos utilizados já apresentarão uma deficiência de representação logo nos dados de entrada, o que gera um segundo problema, que é a ineficiência dos sistemas estatísticos aplicados em cada país (BAUCHET; MORDUCH, 2009). Além disso, as instituições financeiras usualmente utilizam mecanismos tecnológicos para oferecer e controlar seus empréstimos.

Para tornar a aplicação de ML mais clara, o presente estudo se utiliza do site de financiamento coletivo Kiva.org para exemplificação. Conforme explicitado no tópico 2.1.2, a instituição é baseada em um modelo de financiamento coletivo, na qual os empreendedores (ou ONGS locais parceiras do Kiva) criam anúncios de seus negócios no site, podendo ser financiados por pessoas ao redor do mundo.

Nos tópicos abaixo, é feita uma análise de como o viés pode ocorrer no microcrédito através do site Kiva.org, de acordo com as classificações exibidas acima, e nos modelos de análise de saídas citados no tópico 2.2.4. Além disso, serão propostas possíveis soluções para os problemas relacionados ao viés, baseados na literatura.

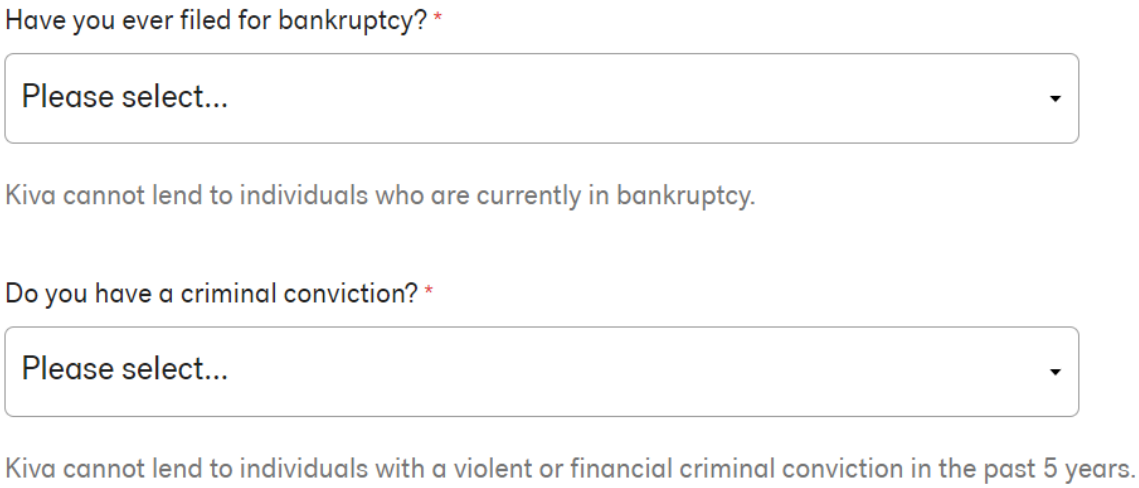
Vale ressaltar que as percepções e classificações dos vieses do site são baseadas na bibliografia já publicada sobre o assunto e na percepção da autora, uma vez que não se teve acesso ao código completo do Kiva como fonte. A plataforma apresenta comunidades *open source*, porém não disponibiliza seu código para estudos mais aprofundados (ANEXO B).

3.3.1 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE CLASSIFICAÇÃO

Em primeiro momento, serão avaliadas as possibilidades de viés nos problemas de classificação. Por se tratarem de problemas que trazem saídas baseadas em pré-classificações do código, o método de aprendizado supervisionado é o que apresenta maiores chances de se enquadrar nessa categoria.

Para que o microempreendedor possa anunciar seu negócio no site, o mesmo precisa preencher um questionário com suas informações pessoais. Diante das informações que são levantadas, pode ocorrer o viés de treinamento, no qual as informações apresentam tendências do programador ou do próprio usuário, que podem interferir nas saídas e escolhas dos credores. Existem algumas categorias que podem tendenciar as escolhas, como as perguntas *"Do you have a criminal conviction?"* e "*Have you ever filed for bankruptcy?*", exemplificadas na Figura 8, que significam, respectivamente: "Você já foi condenado criminalmente?" e "Você já foi à falência?". Perguntas como essas podem fazer com que os anunciantes não sejam escolhidos pelos credores, baseados em informações anteriores e que não representam mais a condição do empreendedor.

Figura 8 – Categorias de cadastro Kiva: condenação e falência



Fonte: Kiva.org

Já no caso da classificação que o código realiza com base nas informações inseridas, sem utilizar a auto classificação de acordo com o formulário de cadastro, pode ocorrer também o viés de treinamento, no qual as *tags* são indevidas, como um caso de anúncio de campanha de apoio a confecção que foi classificada como apoio para combate à fome, e a *tag* "*Women*" (mulher, em inglês), que deixa grupos vulneráveis exclusos, como pessoas trans. Ainda sobre as *tags* referentes aos gêneros, pode-se perceber mais uma possibilidade de tendenciamento, que ocorre quando o usuário seleciona uma categoria diferente de "*Women"*, como

“Agricultura”, que permite no passo seguinte a utilização de um filtro de gênero. No filtro, é exibida a opção “não-binário”, o que não é possível ao selecionar inicialmente a categoria *Women*.

Para o viés de processamento, pode ser citada a possibilidade de manipulação de dados através das habilidades do programador, como desconsiderar informações importantes e alteração das saídas para privilegiar certos grupos. Os próprios modelos matemáticos e estatísticos passíveis de aplicação podem se diferir expressivamente quanto às saídas, também podendo privilegiar certos grupos inconscientemente.

No viés de foco, podem ser citadas categorias que não são tão relevantes para sociedades em específico. Um exemplo é a pergunta *“My Business is legal at the federal level (e.g. no cannabis products. Due to PayPal policy, we cannot support loans for CBD or Hemp products.)”*(Meu negócio é legal no nível federal (por exemplo, nenhum produto de cannabis. (“Devido à política do PayPal, não podemos oferecer empréstimos para produtos de CBD ou derivados de cannabis”, do ingês), na qual o empreendedor responde se seu negócio é legal com base nas leis. Nesse caso, em países que a legislação é mais branda, as campanhas ficam limitadas, excluindo possibilidades para anúncios ligados à cannabis, tabaco, penhoras, armas, pirâmide e entretenimento adulto; refinanciamento de dívidas; compra de ações. Ainda com o exemplo da pergunta em questão, pode ocorrer o viés de contexto, pois essa informação pode ser relevante em um país mais “rígido” e irrelevante em outras localidades.

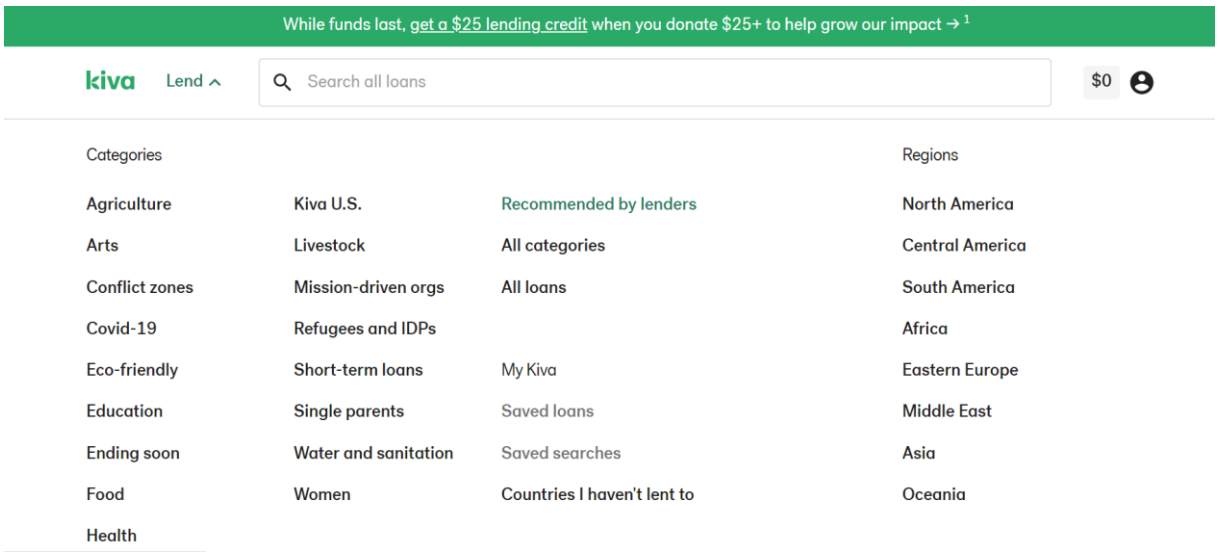
Outro exemplo de viés de contexto são *tags* que associam as campanhas à vulnerabilidade, fome e pobreza. Os níveis de pobreza variam de acordo com as condições de cada país, sendo um país com PIB mais alto com o corte de pobreza maior do que países mais pobres.

Conforme supracitado, todos os anúncios do site são enquadrados em alguma categoria, seja ela de país, gênero, modelo de negócio, etc; sendo elas baseadas nas informações que são fornecidas pelo microempreendedor, distribuídas dentro de grupos pré-estabelecidos pelo programador ou através da autoclassificação. Ambas as possibilidades estão sujeitas ao viés.

A Figura 9 é uma captura de tela do site, na qual algumas das categorias existentes no site são exibidas.

Ainda ao observar a Figura 9, é possível perceber a existência de dois filtros para empreendimentos dos Estados Unidos, um como a categoria “Kiva U.S” e outro na escolha da região América do Norte (*North America*). Isso faz com que as chances de doação para tal país sejam maiores.

Figura 9 – Categorias do Site Kiva.org



Fonte: Kiva.org

Da mesma forma, algumas categorias se sobrepõem, possibilitando a maior possibilidade de seleção de projetos classificados em mais de uma categoria. Ou seja, caso o usuário deseje aparecer em pesquisas de todas as categorias do site, será possível, já que as mesmas não são mutuamente excludentes. Esse seria mais um exemplo de viés de treinamento.

Todavia, podemos perceber que a instituição Kiva.org já realiza esforços para diminuir as possibilidades de viés, como a inclusão da categoria "*Countries I haven´t lent to*" (Países para os quais ainda não emprestei), que evita a exclusão de diferentes etnias nos filtros. Para que esse efeito se espalhasse no site, poderiam ser inclusas opções como essa para todas as categorias.

No entanto, vale ressaltar que a possibilidade do viés continuará presente nos códigos mesmo após as mudanças acima, visto que cada indivíduo tem suas peculiaridades e enxerga a realidade e as categorias de uma maneira, o que de alguma forma estará enviesando os dados de entrada e a classificação do algoritmo.

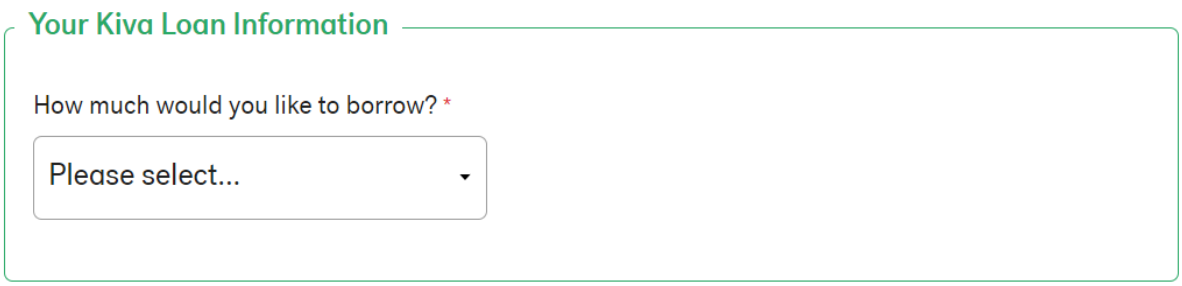
3.3.2 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE REGRESSÃO

Nos casos de regressão, os códigos apresentam saídas numéricas baseadas nas características dos dados históricos, exibindo quais entradas se comportariam similarmente.

Sendo assim, pode ocorrer o viés de processamento, relacionado a quanto o usuário pode pegar emprestado. No questionário inicial, existem perguntas como *How uch would you like to borrow?* (O quanto você gostaria de pegar emprestado) (a pergunta apresenta uma lista

suspensa com valores de $1.000 a $15.000). A Figura 10 representa a pergunta no formulário inicial. Sugere-se que, através dessa informação, o código pode acabar limitando o quanto a pessoa pode pegar emprestado ao comparar perfis que apresentem características semelhantes, mas que não necessariamente apresentam o mesmo comportamento no quesito pagamento do empréstimo.

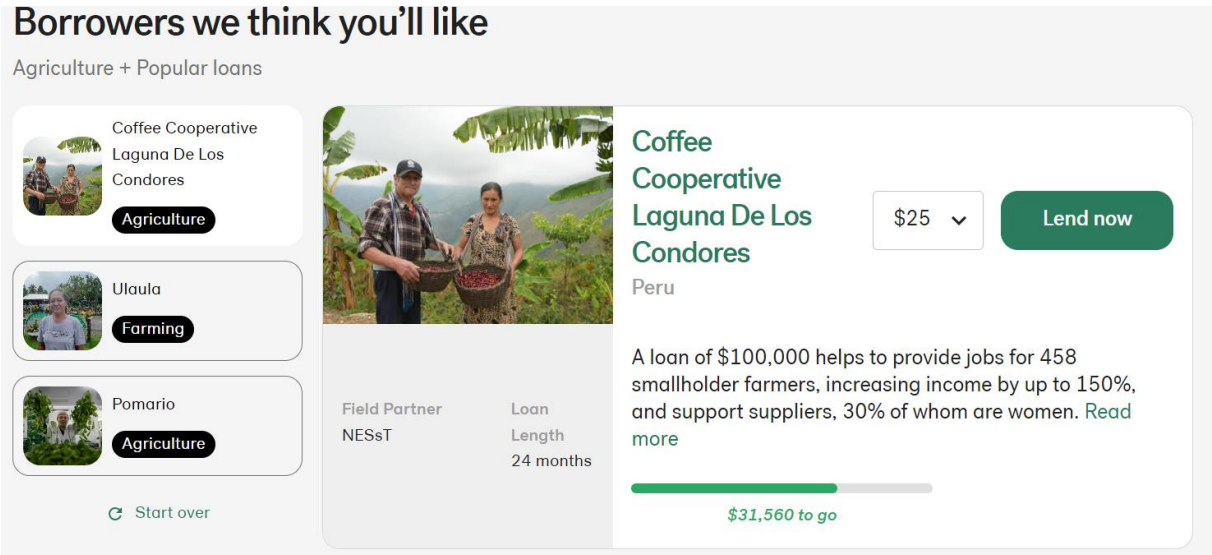
Figura 10 - Categorias de cadastro Kiva: valor do empréstimo



Fonte: Kiva.org

Além do caso de limitação da quantia, também pode ocorrer o viés de processamento com o filtro tendências (*Popular Loans*), que aparece em todas as categorias, conforme pode ser observado na Figura 11, que é um exemplo da categoria nas opções de anúncio relacionados à agricultura.

Figura 11 – Categoria de tendência



Fonte: Kiva.org

O filtro é baseado na velocidade com que as doações caem, que ajudam a inferir qual campanha está mais próxima de acabar mais rápido, o que depende do ferramental estatístico e matemático escolhido, além do fato de que o tempo de duração das campanhas também depende do valor pedido e dos montantes doados. Campanhas de valores menores com velocidade de doação menor podem acabar antes de campanhas de maior valor e com velocidades de doação maiores. Sendo assim, um modelo matemático pode privilegiar um anúncio em detrimento de outros. Vale citar o viés de contexto, no qual os dados históricos são baseados em uma sociedade e as saídas são aplicadas em outra, que apresenta características diferentes. Um exemplo é criar uma base histórica com dados de empreendedores da Ásia e inserir informações da Índia para serem estimados.

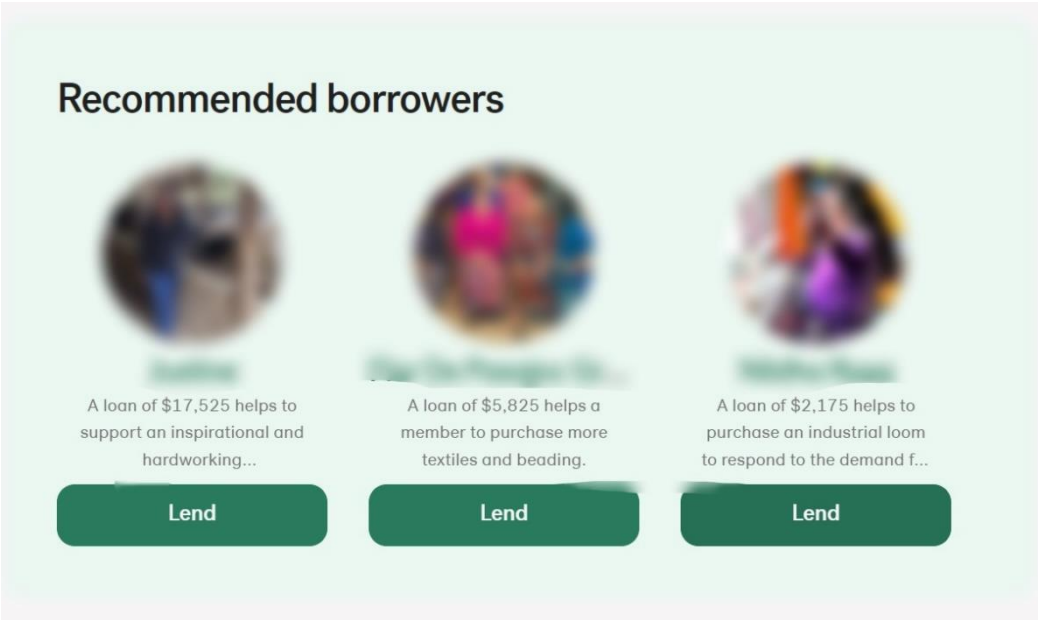
3.3.3 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE AGRUPAMENTOS

O terceiro grupo de problemas a ser abordado é o de agrupamentos, que categorizam os dados dentro de grupos, com base naqueles com características similares entre si e que mais se diferem entre os grupos.

À vista disso, surge a possibilidade de viés de treinamento, no qual o código clusteriza de acordo com os anúncios mais recomendados pelos usuários, conforme a Figura 12. Com isso, os empreendedores que ainda não receberam muitas doações seguem deixados de lado, o que favorece os anúncios já em tendência. Vale citar também outra forma de viés de treinamento, que ocorre na formação de empréstimos com perfis semelhantes. Não se sabe quais são os critérios, mas espera-se que estejam relacionados às categorias de cada anúncio, a média de dias para atingir a meta, pagamento do empréstimo, etc. Com isso, novamente alguns anúncios podem ser priorizados em detrimento de outros. A clusterização com base na formação de grupos de anúncios semelhantes também pode representar um viés de processamento, pois depende do modelo matemático que é escolhido.

Por fim, para os casos de agrupamento, temos o viés de contexto, no qual os anúncios considerados como “legais” interferem nas escolhas dos credores pois cada país apresenta uma legislação diferente. Muitos países diferem em legislação quanto ao cultivo de cannabis, por exemplo, e anúncios relacionados ao tema não são permitidos pois o Kiva é baseado no código legal dos Estados Unidos, na qual cada estado apresenta um posicionamento quanto ao uso da maconha, seja ele recreativo ou medicinal.

Figura 12 – Recomendações



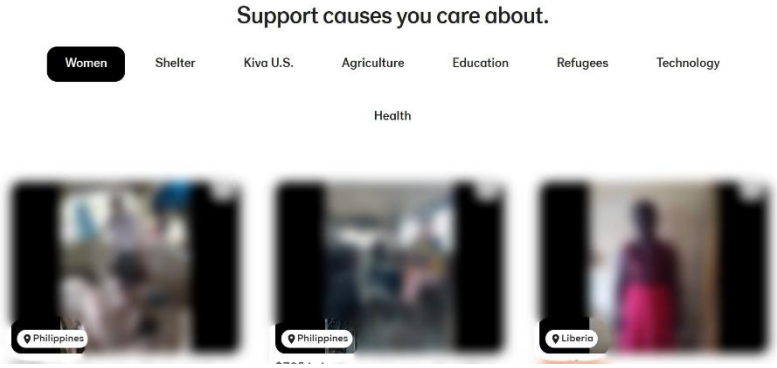
Fonte: Kiva.org

3.3.4 MICROCRÉDITO E O VIÉS NOS PROBLEMAS DE ASSOCIAÇÃO

Para encerrar o capítulo das possibilidades de vieses nos problemas de análise das saídas, serão avaliados os problemas de associação.

Um dos principais exemplos de associação é a barra de pesquisa do site Kiva.org. Mesmo não tendo o código do site em mãos para confirmar o modelo utilizado, é sabido que o mecanismo mais comum em recomendações é verificar quais os interesses do usuário tendo como base escolhas prévias, gerando um perfil de preferências, no qual a máquina aprende a sugerir elementos com características comuns e recorrentes. A Figura 13 é um exemplo de recomendações feitas no site.

Figura 13 – Recomendações do site Kiva



Fonte: Kiva.org

Sistemas de recomendação são utilizados em diversos ramos do cotidiano da população, como em notícias, plataformas de *streaming*, plataformas de música, etc. O modelo verifica quais são as preferências do usuário através de suas escolhas anteriores, e oferece recomendações baseadas nos interesses, preferências e necessidades do mesmo (LIU; BURKE, 2018). O mesmo ocorre na plataforma Kiva.org, que apresenta novos anúncios a partir das escolhas anteriores do usuário.

Pode ocorrer o viés de treinamento, no qual o código deixa em segundo plano os anúncios de empreendedores que se enquadram nos quesitos de falências e condenações criminais, pois esses representam uma baixa aceitação dos consumidores do site.

O viés de processamento pode ocorrer quando o mecanismo de recomendação indica perfis baseados nas escolhas anteriores, sendo mais propenso a indicar somente empreendimentos com características semelhantes, colocando em segundo plano a possibilidade de diversificação e inclusão, sem falar no fato de as saídas dependerem do ferramental estatístico escolhido pelo programador. O viés de processamento também pode ser observado na frequência de acesso à plataforma, que identifica os usuários que estão mais propensos a doar e apresentar mais recomendações para estes, o que também vai depender de qual o modelo estatístico foi escolhido para avaliação.

Já o viés de foco pode ser evidenciado na opção convidar amigos, o qual pode tendenciar as recomendações do novo usuário com base nas preferências de quem indicou.

Por fim, temos o viés de contexto, que pode ser exemplificado no fato de haver mais recomendados asiáticos pelo simples fato de haver mais usuários asiáticos. Tal fato pode ser observado através do estudo realizado por Liu e Burke em 2018.

O estudo de Liu e Burke (2018) analisa o mecanismo de ranqueamento de mercados online, e propõe um modelo que balanceie a qualidade das sugestões personalizadas, dando chance a todos o micro negócios que são expostos no site. Os autores elaboram um algoritmo que percebe quais usuários são mais tolerantes às diferenças, e permite que anúncios de fornecedores diferentes das últimas escolhas do usuário apareçam. Um dos experimentos realizados no estudo é sobre o site Kiva.org, no qual são avaliadas as transações no período de oito meses.

Como pode ser observado na figura 14, durante os oito meses analisados, os empréstimos foram realizados para diversas regiões. A primeira coluna apresenta a região, a segunda a quantidade de empréstimos e a terceira a porcentagem do total.

Figura 14 – Total de empréstimos do site Kiva.org por região

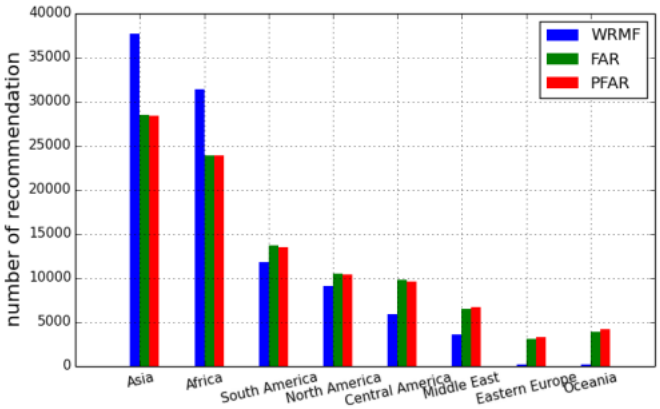
Table 3: Inventory of each region in Kiva.org.

Region	Inventory	% of Total
Asia	1619	25.3%
Africa	1950	30.4%
South America	1146	17.9%
North America	389	6.1%
Central America	629	9.8%
Middle East	383	6.0%
Eastern Europe	166	2.6%
Oceania	128	2.0%

Fonte: LIU; BURKE, 2018

Já na figura 15, é possível observar a distribuição das recomendações por país. Verifica-se, através das colunas em azul, como os anúncios são desiguais, tendo a Ásia um alto número de recomendações, e a Oceania quase nenhuma.

Figura 15 – Total de recomendações por região[6]



Fonte: LIU; BURKE, 2018

Os autores percebem que o site não oferece uma lista de sugestões vasta, então aplicam o modelo de re-ranqueamento, inserindo pseudo características para cada perfil de usuário (seja ele anunciante ou potencial fonte de empréstimo), o que gera uma maior lista de recomendações para o site. A coluna laranja apresenta as novas saídas de recomendações, após as mudanças de recomendações realizadas com base no algoritmo que foi elaborado pelos autores. Através dela é percebido que foram feitas reduções nas recomendações de empreendedores da Ásia e África, o que consequentemente resultou em um aumento nos continentes da Europa Oriental e Ásia.

[6] *FAR - Fairness-Aware Re-ranking*
PFAR - Personalized Fairness-Aware Re-ranking

O estudo de caso de Liu e Burke (2018) é um exemplo que como o viés de processamento algorítmico é aplicado no Kiva.org, e propõe uma saída para que esse problema seja mitigado. Os autores mostram que o site utiliza características muito básicas para realizar as recomendações, e indicam a utilização de classes mais vastas, que representam ainda mais os interesses do usuário, para que as recomendações fiquem mais justas para a sociedade.

4. CONCLUSÕES

A partir dos exemplos e estudos supracitados, fica evidente a importância do aprendizado de máquina no dia a dia das sociedades. Com os avanços na tecnologia, o volume de dados que são gerados e precisam ser processados aumenta de maneira exponencial, o que torna quase impossível a não utilização de *Machine Learning*.

Em contrapartida, é possível perceber que as técnicas de ML não estão isentas de falhas e lacunas, como é o caso do viés algorítmico. Os problemas surgem desde a concepção do código, até a interpretação de suas saídas. Muitas são as possibilidades de enviesamento e réplica de tendências de uma sociedade desigual. Para isso, é importante que as empresas e usuários fiquem atentos aos possíveis pré-conceitos que podem ser perpetuados.

Dando foco à interseção entre as técnicas de ML e o empreendedorismo social, foi realizado um recorte na política de microcrédito, tendo como base a instituição Kiva.org. Como o presente trabalho não teve acesso ao código completo do site, foram feitas algumas suposições e interpretações sobre pesquisas em materiais já publicados, para demonstrar como o viés pode interferir na concessão de crédito. Diante disso, foi possível perceber que a instituição pode, e deve, utilizar ML para aprimorar sua atuação e efetividade. Porém, deve estar atenta às suas falhas, e buscar sempre sua mitigação.

Um dos exemplos avaliados é a classificação dos usuários, que pode ser baseada em moldes pré definidos ou então na auto classificação. Para evitar excluir minorias, o site pode criar categorias mais gerais, e também permitir que usuários indiquem novas categorias. O aprendizado deve ser diário, e deve refletir as particularidades da sociedade.

Após a análise de todas as possibilidades de viés nos métodos de aprendizagem e nos modelos de análise das saídas, foi elaborado um quadro com um resumo de todas as percepções, apresentado na Tabela 4. Além disso, foi feito um resumo com as principais características dos tipos de vieses (Tabela 2) e com os métodos de aprendizado analisados (Tabela 3).

Tabela 2 – Resumo Vieses

VIÉS DE TREINAMENTO	As informações de entrada ou de treinamento fornecidas representam uma tendência do programador ou do usuário, transmitindo seus preconceitos e preferências ao modelo.
VIÉS DE PROCESSAMENTO	Programador utiliza seu conhecimento e as ferramentas de ML para mascarar resultados e alterar as saídas, manipulando os dados que serão repassados para o público.

VIÉS DE FOCO	Utilizadas categorias de classificação irrelevantes, deixando de lado informações estatisticamente valiosas, alterando as saídas do programa. Pode ocorrer também de existirem categorias ilegais, que não podem ser utilizadas para análise.
VIÉS DE CONTEXTO	Inconsistências entre o código e o contexto no qual o mesmo é aplicado.
VIÉS DE INTERPRETAÇÃO	O usuário apresenta conceitos e valores que diferem do código e de suas saídas. Nesse caso, a informação trazida pelo algoritmo será interpretada de maneira enviesada, podendo ser distorcida por quem a analisa.

Fonte: Elaborado pela autora

Tabela 3 – Resumo: Métodos de análise de saídas

Classificação	Aprender uma regra que entregue como saída um mapeamento correto das entradas dentro de classes pré definidas. Nesse tipo de problema, o objetivo é sempre categorizar.
Regressão	Avaliam dados históricos e retornam quais dados de entradas vão apresentar características similares. Apresenta saídas numéricas
Agrupamento	Categorizam os dados dentro de subconjuntos (grupos). São problemas que visam agrupar elementos com propriedades em comum, que os aproximem entre si e os diferenciem de outros grupos.
Associação	Avaliar a primeira escolha do usuário e prever a segunda.

Fonte: Elaborado pela autora

Tabela 4 – Quadro resumo da classificação dos vieses

		APRENDIZAGEM SUPERVISIONADA	APRENDIZAGEM NÃO SUPERVISIONADA OU SEMI SUPERVISIONADA	APRENDIZADO POR REFORÇO
CLASSIFICAÇÃO	VIÉS DE TREINAMENTO	* Informações apresentam tendências do programador ou do próprio usuário, que podem interferir nas saídas e escolhas dos credores. Ex: Campos “se já faliu” ou “se já foi condenado criminalmente”	* Tags erradas, por exemplo: - Campanha para apoio à confecção classificada como para fome - Tag “Women” (mulher, em inglês), que deixa grupos vulneráveis exclusos, como pessoas trans	
	VIÉS DE PROCESSAMENTO	Desconsiderar informações importantes e alteração das saídas para privilegiar certos grupos		
	VIÉS DE FOCO	Limitam as campanhas proibitivas às ligadas à cannabis, tabaco, penhoras, armas, pirâmide e entretenimento adulto; refinanciamento de dívidas;		
	VIÉS DE CONTEXTO	Pesquisas realizadas em um grupo e aplicadas para um código utilizado em outra amostra da sociedade. Como é o caso do filtro de negócio legal ou não	Palavras-chave são “livres”: insegurança alimentar, sustentabilidade, etc., podem ter sentidos diferentes em diferentes contextos.	
REGRESSÃO	VIÉS DE TREINAMENTO		Filtro “Tendências”. Sugere que baseado na velocidade com que as doações	

			caem, que ajudam a inferir qual campanha está mais próxima de acabar mais rápido.	
	VIÉS DE PROCESSAMENTO	Verificar o quanto cada usuário pode pegar emprestado, baseado em outros perfis com características similares.	Filtro "Tendências". Sugere que baseado na velocidade com que as doações caem, que ajudam a inferir qual campanha está mais próxima de acabar mais rápido - depende do ferramental estatístico e matemático escolhido.	
	VIÉS DE FOCO			
	VIÉS DE CONTEXTO	Utilizar a base de dados de um local e aplicar as saídas em local com condições diferentes		
AGRUPAMENTOS	VIÉS DE TREINAMENTO	Mais recomendados por outros usuários	Formação das equipes de empréstimos com perfis semelhantes	
	VIÉS DE PROCESSAMENTO		Formação das equipes de empréstimos com perfis semelhantes – c– depende do ferramental estatístico e matemático escolhido.	
	VIÉS DE FOCO			
	VIÉS DE CONTEXTO	Países que diferem em legislação quanto ao cultivo de cannabis, por exemplo.		
ASSOCIAÇÃO	VIÉS DE TREINAMENTO	Campos "se já faliu" ou "se já foi condenado criminalmente".		A ferramenta de recomendação por si só é um

	VIÉS DE PROCESSAMENTO	"Quem doa para esse tipo de setor, atividade, território estará mais propenso a doar para esse" – depende do ferramental estatístico e matemático escolhido.	Frequência de acesso à plataforma em razão do histórico de aportes e empréstimos - estimativa de quando estará mais propenso a doar – depende do ferramental estatístico e matemático escolhido.	exemplo do método não supervisionado, podendo apresentar todas as possibilidades citadas nas linhas de Associação
	VIÉS DE FOCO	"Convidar amigos" – pode tendenciar as recomendações do novo usuário com base nas preferências de quem indicou.		
	VIÉS DE CONTEXTO	Mais recomendados asiáticos pelo simples fato de haver mais usuários asiáticos.		

Fonte: Elaborado pela autora

Por fim, vale ressaltar que, mesmo com a intensificação dos mecanismos automatizados, as chances de viés e réplica de preconceitos sempre estará presente, visto que a atuação do ser humano sobre as mesmas também estará. Códigos são elaborados por programadores, que já têm uma visão particular da realidade, opiniões e preferências. Por isso se faz tão importante o estudo de como minimizar tais tendências, para que cada vez menos pessoas sejam prejudicadas com essa situação. Ademais, o presente trabalho, por ser baseado em percepções da autora e de estudos de autorias diversas, também apresenta vieses dos mesmos.

A tecnologia deve ser uma grande aliada na melhoria da qualidade de vida, e não um artifício para afastar cada vez mais as pessoas e criar mais lacunas na sociedade. Para isso, deve ser intensificada a pesquisa e a avaliação dos modelos já utilizados e de novas possibilidades, unindo conhecimentos multidisciplinares, como engenharia, economia, tecnologia da informação, medicina, etc. Quanto mais próximo da realidade forem as entradas e os parâmetros

que são utilizados no ambiente de máquina, maiores são as chances de conseguir atender suas demandas e suprir necessidades que não são atendidas nos dias atuais.

REFERÊNCIAS

ABEPRO. **Tabela de áreas da Engenharia de Produção**. Elaborado pela Comissão de Graduação e referendado no GT de Graduação do Encep 08 Enegep 08 – 2008. Disponível em http://portal.abepro.org.br/enegep/wp-content/uploads/2021/03/Areas-da-Engenharia-de-Producao_ENEGEP-2021.pdf. Acesso em: 19 jun. 2022.

ALVES, Marco Antônio Sousa; ANDRADE, Otávio Morato de. Da "caixa-preta" à "caixa de vidro": o uso da explainable artificial intelligence (XAI) para reduzir a opacidade e enfrentar o enviesamento em modelos algorítmicos. RDP, Brasília, v. 18, n. 100, p. 349-373, Out-Dez 2021. DOI 10.11117/rdp.v18i100.5973. Disponível em: https://www.researchgate.net/publication/359178400_Da_caixa-preta_a_caixa_de_vidro_o_uso_da_explainable_artificial_intelligence_XAI_para_reduzir_a_opacidade_e_enfrentar_o_enviesamento_em_modelos_algoritmicos. Acesso em: 25 set. 2022.

ÁVILA, L. V.; DA ROCHA, M. P.; ARIGONY, M. M.; DILL, R. A.; MAZZA, V. M. D. Souza. **NEGÓCIOS COM IMPACTO SOCIAL: CARACTERÍSTICAS, MODELOS E MÉTRICAS DE AVALIÇÃO**. Gestão e Desenvolvimento em Revista, *[S. l.]*, v. 2, n. 1, p. p. 4–13, 2016. DOI: 10.48075/gdemrevista.v2i1.14754. Disponível em: https://saber.unioeste.br/index.php/gestaoedesenvolvimento/article/view/14754. Acesso em: 20 jun. 2022.

BAUCHET, Jonathan; MORDUCH, Jonathan. **Selective knowledge: Reporting biases in microfinance data**. Finance Access Initiative : Impact Through Research: NYU Wagner Graduate School, New York, 8 jun. 2009. Disponível em: https://wagner.nyu.edu/files/faculty/publications/2008-06-Selective-knowledge.pdf. Acesso em: 14 dez. 2022.

BAZZAN, A. L. C. **Contribuições de aprendizado por reforço em escolha de rota e controle semafórico.** Estudos Avançados, [S. l.], v. 35, n. 101, p. 95-110, 2021. DOI: 10.1590/s0103-4014.2021.35101.008. Disponível em: https://www.revistas.usp.br/eav/article/view/185038. Acesso em: 4 dez. 2022.

BOSE, Monica. **Empreendedorismo social e promoção do desenvolvimento local**. 2012. 198 p. Tese (Doutorado Administração e Contabilidade) - Universidade de São Paulo, São Paulo, 2012. Disponível em: https://doc-0s-74-apps-viewer.googleusercontent.com/viewer/secure/pdf/ksdrv3m3imckfvbb76hbb1a3s1p006qn/6fd4qnksg8gpp6u8vmoum0ppad1eg8qc/1656292800000/drive/12315622418302536202/ACFrOgAuPtGFS5-bw_ehZRyj_YU6CBMI7I4S-HBBsYS1Iky_VJBNMBkRy4tIRxJYirI7R1ieChbPw9WdhsfkpmZr9saPjPIdJNspulT3Eha6skM6YLf86-bimsK9tUAshMrAymLpNGh4vo1gxX43?print=true. Acesso em: 27 jun. 2022.

BORNSTEIN, D. **How to change the world: Social entrepreneurship and the power of ideas**. Oxford: Oxford University Press, 2007.

BRANDÃO, Daniel; ARIDA, Anna Livia; CRUZ, Celia . **Métricas em negócios de impacto social: FUNDAMENTOS**. Move Social e Instituto de Cidadania Empresarial (ICE), [*S. l.*], p. 1-15, 2014. Disponível em: https://sinapse.gife.org.br/download/metricas-em-negocios-de-impacto-social-

fundamentos#:~:text=O%20artigo%20aborda%20as%20referencias,explorar%20as%20ferramentas%20centrais%20existentes. Acesso em: 24 jun. 2022.

BRITO, Jessica de Lima; CARVALHO, Renata Oliveira de. **A importância do empreendedorismo social no desenvolvimento local: um estudo de caso enactus-ufscar.** In: WORKSHOP DE INOVAÇÃO, PESQUISA, ENSINO E EXTENSÃO, 3., 2018, São Carlos, SP. Anais... São Carlos, SP: IFSP, 2018. p. 123-126. ISSN 2525-9377. Acesso em: 27 jun. 2022.

BRONSTRUP, Felipe César; SIMÕES, Wagner. **EXPLORAÇÃO DE OPORTUNIDADES DE ATUAÇÃO POR ENGENHEIROS DE PRODUÇÃO COMO CIENTISTAS DE DADOS.** Expo Ulbra 2017: XXII Salão de Iniciação Científica e Tecnológica, [s. l.], 2017. Disponível em: http://www.conferencias.ulbra.br/index.php/sic/sic23/paper/viewFile/9217/4388. Acesso em: 18 jul. 2022.

BUTLER, Judith. **Capitalism Has its Limits: Judith Butler discuss the COVID-19 pandemic, and its escalating political and social effects in America**. Verso, [*S. l.*], 30 mar. 2020. Disponível em: https://www.versobooks.com/blogs/4603-capitalism-has-its-limits. Acesso em: 4 jun. 2022.

CAVIQUE, Luís. **Big Data e Data Science**. Boletim da APDIO : IO EM ACÇÃO, [s. l.], n. 51, p. 11-14, 2014. Disponível em: http://hdl.handle.net/10400.2/3918. Acesso em: 17 jul. 2022.

CERNEV, Adrian Kemmer; PROENÇA, Bruna Auad. **MUMBUCA: A PRIMEIRA MOEDA SOCIAL DIGITAL DO BRASIL**. GV Casos: Rev. Bras. de Casos de Ensino em Administração, [s. l.], v. 6, n. 2, 13 dez. 2016. DOI https://doi.org/10.12660/gvcasosv6n2c15. Disponível em: https://bibliotecadigital.fgv.br/ojs/index.php/gvcasos/article/view/61805. Acesso em: 26 maio 2022.

CERNEV, Adrian Kemmer. **MUMBUCA É DINHEIRO**. GV Casos, [s. l.], v. 9, n. 2, 2019. DOI http://dx.doi.org/10.12660/gvcasosv9n2c10. Disponível em: https://bibliotecadigital.fgv.br/ojs/index.php/gvcasos/article/view/80878. Acesso em: 19 jul. 2022.

CHIAVEGATTO FILHO, Alexandre Dias Porto. **Uso de big data em saúde no Brasil: perspectivas para um futuro próximo**. Epidemiol. Serv. Saúde, Brasilia, n. 2, ed. 24, p. 325-332, 2015. DOI 10.5123/S1679-49742015000200015. Disponível em: https://www.scielo.br/j/ress/a/gdPPJMW7YcfK5pk56MJMZPb/?lang=pt&format=pdf. Acesso em: 18 jul. 2022.

CHOO, Jaegul; LEE, Changhyun; LEE, Daniel; ZHA, Hongyuan; PARK, Haesun. **Understanding and promoting micro-finance activities in Kiva.org**. 7th ACM international conference on Web search and data mining, [s. l.], p. 583-592, 2014. DOI https://doi.org/10.1145/2556195.2556253. Disponível em: https://dl.acm.org/doi/abs/10.1145/2556195.2556253?casa_token=P-DshD1uwFAAAAAA:yfmQLA5acf2axliInAx2UWN4NPc2oLs37oA-elXZJLOhx915wnSQp9fgxOcfzoB2HT01xssqXyI. Acesso em: 2 nov. 2022.

CURTY, Renata Gonçalves; SERAFIM, Jucenir da Silva. **A formação em ciência de dados: uma análise preliminar do panorama estadunidense**. Informação & Informação, [S.l.], v. 21, n. 2, p. 307–331, dez. 2016. ISSN 1981-8920. DOI:http://dx.doi.org/10.5433/1981-8920.2016v21n2p307. Disponível em: https://www.uel.br/revistas/uel/index.php/informacao/article/view/27945. Acesso em: 17 jul. 2022.

CRUZ, Guillermo. **AS DUAS FACES DO EMPREENDEDORISMO SOCIAL**. RAUnP, [*s. l.*], v. 5, n. 1, p. 9-20, 20 dez. 2012. DOI https://doi.org/10.21714/raunp.v5i1.324. Disponível em: https://repositorio.unp.br/index.php/raunp/article/view/324. Acesso em: 15 jun. 2022.

DAMBROS, Henrique Vargas. **Predição de alvos de microRNAs utilizando aprendizado semi-supervisionado e classificação baseada em uma única classe**. 2019. 64 p. Monografia (Bacharel em Ciência da Computação) - UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, [S. l.], 2019. Disponível em: https://lume.ufrgs.br/handle/10183/212889. Acesso em: 4 dez. 2022.

DA ROCHA, C. J.; PORTO, L. V.; ABAURRE, H. E. **Discriminação algorítmica no trabalho digital**. Revista de Direitos Humanos e Desenvolvimento Social, [S. l.], v. 1, p. 1–21, 2020. DOI: 10.24220/2675-9160v1e2020a5201. Disponível em: https://periodicos.puc-campinas.edu.br/direitoshumanos/article/view/5201. Acesso em: 6 nov. 2022.

DA SILVA, Maria de Fátima; MOURA, Laysce Rocha; JUNQUEIRA, Luciano Antonio Prates. **AS INTERFACES ENTRE EMPREENDEDORISMO SOCIAL, NEGÓCIOS SOCIAIS E REDES SOCIAIS NO CAMPO SOCIAL**. Revista de Ciências da Administração, [*s. l.*], v. 17, n. 42, p. 121-130, 19 ago. 2015. DOI http://dx.doi.org/10.5007/2175-8077.2015v17n42p121. Disponível em: https://periodicos.ufsc.br/index.php/adm/article/view/2175-8077.2015v17n42p121/pdf_67. Acesso em: 27 jun. 2022.

DANKS, David; LONDON, Alex John. **Algorithmic Bias in Autonomous Systems**. IJCAI, [s. l.], p. 4691-4697, 2017. Disponível em: https://doity.com.br/media/doity/submissoes/artigo-c1c98eb6f48bf568311d987ec08a85d28375b9bd-segundo_arquivo.pdf. Acesso em: 5 nov. 2022.

DE CARVALHO, Allan Pereira. **Viés Algorítmico e Discriminação: Possíveis soluções regulatórias para o Brasil**. 2020. 77 p. Trabalho de Conclusão de Curso (Bacharelado em Direito) - Universidade Federal do Rio Grande do Sul, [S. l.], 2020. DOI http://hdl.handle.net/10183/221914. Disponível em: https://www.lume.ufrgs.br/handle/10183/221914. Acesso em: 25 set. 2022.

DE OLIVEIRA, Luiz Henrique Rodrigues; DO RÊGO, Thais Gaudêncio; DINIZ, Josedilton Alves. **Previsão de Valores de Aquisições Governamentais: o Uso dos Conceitos de Data Science e Machine Learning**. XVI Congresso USP de Iniciação Científica em Contabilidade, São Paulo, 2019. Disponível em: https://congressousp.fipecafi.org/anais/19UspInternational/ArtigosDownload/1514.pdf. Acesso em: 18 jul. 2022.

DOLABELA, F. Oficina do Empreendedor. Rio de Janeiro: Sexante, 2008.

DOS PASSOS, Danielle Sandler. **BIG DATA, DATA SCIENCE E SEUS CONTRIBUTOS PARA O AVANÇO NO USO DA OPEN SOURCE INTELLIGENCE**. Revista Eletrônica Sistemas & Gestão, [s. l.], v. 11, n. 4, p. 392-396, 2016. DOI 10.20985/1980-5160.2016.v11n4.1026. Disponível em: https://revistasg.uff.br/sg/article/view/1026/524. Acesso em: 17 jul. 2022.

ESCOVEDO, Tatiana; KOSHIYAMA, Adriano. Introdução a Data Science: Algoritmos de Machine Learning e métodos de análise. [S. l.]: Casa do Código, 2020. 338 p.

ESTY, Katharine. **Lessons from Muhammad Yunus and the Grameen Bank**. Od Practitioner, v. 43, n. 1, p. 24-28, 2011.Disponível em: https://d1wqtxts1xzle7.cloudfront.net/52599363/OD_Practitioner_Proof_Rev5-with-cover-page-v2.pdf?Expires=1654395802&Signature=JvIGFxUqo5yPqwPlEHCkmn77~Prbsw-weybcWNmpQd3WUcX2VocvY~4oTeCcjhuEd7mdnCi2WoauUk9q18DgT0vO4OVTYHFBHjxCPkF4lQHJFEnE89vGNycH6kQJ6YaROPZgoPEee8VgjAptRLEOMUdUZkJ2fVJb-Ng7OMLiohVuQRDSnMd9gL~bvOGlrinn~rADWK9OsS4CNlrPm-8Fc-CQ8UukpgYmDIQ3k8HlWtGo22pXfSBC1U8J-N9oHmMWbRXANi5ukEW6Mb6ER74e~J-wMRHVc1kStYSTd1ieSGRHdSYm9HMihVeYiat8qz~n8m8JSwwIjWYSxd-NIRZwVw__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=26. Acesso em: 4 jun. 2022.

FAGUNDES, Flávio; ZEIFERT, Anna Paula Bagetti; SIQUEIRA, Carelisa Stoffel de. **REDUZIR A DESIGUALDADE DENTRO DOS PAÍSES E ENTRE ELES: REFLETINDO SOBRE O ODS 10 DA AGENDA 2030 E A JUSTIÇA SOCIAL EM TEMPOS DE RACIONALIDADE NEOLIBERAL**. Salão do Conhecimento: UNIJUÍ 2020, [*s. l.*], v. 6, n. 6, 21 out. 2020. Disponível em: https://publicacoeseventos.unijui.edu.br/index.php/salaoconhecimento/article/view/18592. Acesso em: 4 jun. 2022.

FEITOSA, Wilian Ramalho; DE MOURA, Mauricio José Serpa Barros; BARBOSA, Robson; ROSSI, Fernando Luis. **O INSTRUMENTO DO MICROCRÉDITO: LEALDADE DO CLIENTE E IMPACTOS SOCIAIS**. Revista de Geopolítica, [s. l.], v. 13, n. 1, p. 14-27, 2022. Disponível em: http://www.revistageopolitica.com.br/index.php/revistageopolitica/article/view/390. Acesso em: 19 jul. 2022.

FLANNERY, Matt. **Kiva and the Birth of Person-to-Person Microfinance**. v. 2, n. 1-2, p. 31–56, 2007. Disponível em: <https://direct.mit.edu/itgg/article/2/1-2/31/9483/Kiva-and-the-Birth-of-Person-to-Person>. Acesso em: 4 jun. 2023.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 5 ed., Atlas, São Paulo, 2010.

GRANVILLE, Vincent. **Developing Analytic Talent: Becoming a Data Scientist**. 1. ed. [S. l.]: John Wiley & Sons, 2014. 310 p. ISBN 1118810082.

IZBICKI, Rafael; NASCIMENTO, Carla Ferreira do; DUARTE, Yeda Aparecida de Oliveira; FILHO, Alexandre Dias Porto Chiavegatto; SANTOS, Hellen Geremias dos. **Machine**

learning para análises preditivas em saúde: exemplo de aplicação para predizer óbito em idosos de São Paulo, Brasil. Cadernos de Saúde Pública, [s. l.], p. 1-19, 2019. DOI 10.1590/0102-311X00050818. Disponível em: https://www.scielo.br/j/csp/a/jyhKL6G4dZhcbchMD6bcS8s/abstract/?lang=pt. Acesso em: 4 dez. 2022.

JANIESCH, Christian; ZSCHECH, Patrick; HEINRICH, Kai. **Machine learning and deep learning**. Electronic Markets, [s. l.], v. 31, p. 685–695, 8 abr. 2021. DOI https://doi.org/10.1007/s12525-021-00475-2. Disponível em: https://link.springer.com/content/pdf/10.1007/s12525-021-00475-2.pdf. Acesso em: 19 jul. 2022.

JORGE, Ariany Lima; PEREIRA, Erik Ribeiro; OLIVEIRA, Christian Sousa de; FERREIRA, Eduardo dos Santos; MENON, Edmara Toledo Ninzoli; DINIZ, Susana Nogueira; PEZUK, Julia Alejandra. **MicroRNAs: entendendo seu papel como reguladores da expressão gênica e seu envolvimento no câncer**. einstein (São Paulo), São Paulo, v. 19, eRB5996, jul. 2021. https://doi.org/10.31744/einstein_journal/2021RB5996. Acesso em: 02 mai. 2023.

JÚNIOR, Cairo L. Nascimento; WILTGEN, Filipe; SANTOS, Jeeves Lopes dos. **DESENVOLVIMENTO DE UM SISTEMA DE APRENDIZADO PARA O CONTROLE DO CAMINHAR DE UM ROBÔ UTILIZANDO APRENDIZADO POR REFORÇO**. XVIII Congresso Brasileiro de Automática, Bonito, p. 5028-5035, 2010. Disponível em: https://www.researchgate.net/publication/354707795_DESENVOLVIMENTO_DE_UM_SISTEMA_DE_APRENDIZADO_PARA_O_CONTROLE_DO_CAMINHAR_DE_UM_ROBO_UTILIZANDO_APRENDIZADO_POR_REFORCO_DEVELOPMENT_OF_A_LEARNING_SYSTEM_TO_CONTROL_THE_WALKING_OF_A_ROBOT_USING_LEARNING. Acesso em: 4 dez. 2022.

KERLIN, J. A. **Social entorprise a global comparison**. Tufts University Press. 2009.

KIVA.ORG. About Us. *In*: **Kiva.org** : Make a loan, change a life.. [*S. l.*], 2022. Disponível em: https://www.kiva.org/. Acesso em: 4 dez. 2022.

KRITSKI BÁEZ, Saúl Hercán. **POLÍTICAS PÚBLICAS E DESENVOLVIMENTO: ANÁLISE A RESPEITO DA ODS 10**. Anais do EVINCI: UniBrasil,, Curitiba, v. 5, n. 1, p. 116-116, 21 jan. 2020. Disponível em: https://portaldeperiodicos.unibrasil.com.br/index.php/anaisevinci/article/view/4783. Acesso em: 4 jun. 2022.

LATIFEE, Huzzat Islam. **Microcredit and poverty reduction**. In: International Conference on Poverty Reduction through Microcredit. Taksim-Istambul. 2003. Disponível em: https://grameentrust.org/wp-content/uploads/2021/11/3_MICROCREDIT_AND_POVERTY_REDUCTION_JUNE_2003_IN_TURKEYF.pdf . Acesso em: 4 jun. 2022.

LEITE, Joao P.; CARVALHO, Paulo H. P. de; VIEIRA, Robson D. **Modulação e Codificação Adaptativa em Sistemas OFDM Utilizando Aprendizado por Reforço.** XXX SIMPOSIO BRASILEIRO DE TELECOMUNICAÇÕES, Brasília, 16 set. 2012. Disponível em: file:///C:/Users/judy_/Downloads/797.pdf. Acesso em: 20 jul. 2022.

LIU, Weiwen ; BURKE, Robin. **Personalizing Fairness-aware Re-ranking**. FATREC, Canadá, 18 out. 2018. Disponível em: https://arxiv.org/pdf/1809.02921.pdf. Acesso em: 7 maio 2023.

MACHADO, L.C.R.. **Empreendedorismo Social e Inovação Social: enfoque, finalidades e conexões conceituais**. Revista Mosaico, v.11, n.2, p. 35-46, 2020.

MAHESH, Batta. **Machine Learning Algorithms - A Review**. International Journal of Science and Research (IJSR), [s. l.], v. 9, p. 381-386, 1 jan. 2020. DOI 10.21275/ART20203995. Disponível em: https://www.researchgate.net/profile/Batta-Mahesh/publication/344717762_Machine_Learning_Algorithms_-A_Review/links/5f8b2365299bf1b53e2d243a/Machine-Learning-Algorithms-A-Review.pdf?eid=5082902844932096. Acesso em: 19 jul. 2022.

MATSUBARA, Edson Takashi. **O algoritmo de aprendizado semi-supervisionado co-training e sua aplicação na rotulação de documentos**. 2004. 105 p. Dissertação de Mestrado (Ciências de Computação e Matemática Computacional) - USP São Carlos, [S. l.], 2004. Disponível em: https://www.teses.usp.br/teses/disponiveis/55/55134/tde-19082004-092311/pt-br.php. Acesso em: 4 dez. 2022.

NABINGER, Aluísio Moreira. **UTILIZAÇÃO DE ALGORITMOS DO TIPO MACHINE LEARNING SUPERVISIONADO PARA A CARACTERIZAÇÃO DOS RESULTADOS DA COPA DO MUNDO DE FUTEBOL DE 2018**. 2018. 45 p. Trabalho de Conclusão de Curso (Bacharel em Estatística) - UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, Porto Alegre, 2018. Disponível em: https://lume.ufrgs.br/handle/10183/199214?show=full. Acesso em: 4 dez. 2022.

NAIGEBORIN, V. (2010). **Introduçdo ao Universo de Negocios Sociais**. Biblioteca 2.5. In: INSTITUTO NACIONAL DE EMPREENDEDORISMO E INOVAÇÃO. Biblioteca virtual. 2010. Disponível em: < http://inei.org.br/aceleradora25/biblioteca-2.5/negocios-sociais/Introducao_ao_Universo_de_Negocios_Sociais.pdf-.Acesso em 20jun. 2022.

NEGRETTO, Diego Henrique. **Algoritmos de aprendizado semi-supervisionado baseados em grafos aplicados na bioinformática.** 2016. Dissertação (Mestrado em Ciência da Computação) - UNESP, [S. l.], 2016. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/138136/negretto_dh_me_sjrp_int.pdf?sequence=4&isAllowed=y. Acesso em: 20 jul. 2022.

NUÑES, Pedro. **“Desigualdades educativas en tiempos de coronavirus”**, em La Vanguardia, 14 de abril de 2020. Disponível em: http://www.lavanguardiadigital.com.ar/index.php/2020/04/14/desigualdades-educativas-en-tiempos-de-coronavirus/. Acesso em: 4 jun. 2022.

NUNES, Dierle. **INTELIGÊNCIA ARTIFICIAL E DIREITO PROCESSUAL: VIESES ALGORÍTMICOS E OS RISCOS DE ATRIBUIÇÃO DE FUNÇÃO DECISÓRIA ÀS MÁQUINAS.** Revista de Processo, [s. l.], v. 285, p. 421-447, 2018. Disponível em: https://www.academia.edu/37764508/INTELIG%C3%8ANCIA_ARTIFICIAL_E_DIREITO_PROCESSUAL_VIESES_ALGOR%C3%8DTMICOS_E_OS_RISCOS_DE_ATRIBUI%C3%87%C3%83O_DE_FUN%C3%87%C3%83O_DECIS%C3%93RIA_%C3%80S_M%C3%

81QUINAS_Artificial_intelligence_and_procedural_law_algorithmic_bias_and_the_risks_of_assignment_of_decision_making_function_to_machines. Acesso em: 25 set. 2022.

OLIVEIRA, Tatyana Bitencourt Soares de. **Clusterização de dados utilizando técnicas de redes complexas e computação bioinspirada.** 2008. 112 p. Dissertação de Mestrado (Mestrado em Ciências de Computação e Matemática Computacional) - USP - São Carlos, [S. l.], 2008. DOI 10.11606/D.55.2008.tde-01042008-142253. Disponível em: https://www.teses.usp.br/teses/disponiveis/55/55134/tde-01042008-142253/en.php. Acesso em: 25 set. 2022.

ONU, **17 Objetivos para transformar nosso mundo**. 2015. Disponível em: https://nacoesunidas.org/pos2015/. Acesso em: 4 jun. 2022.

OSISANWO , F.Y; AKINSOLA , J.E.T.; AWODELE, O.; HINMIKAIYE , J. O.; OLAKANMI, O.; AKINJOBI , J. **Supervised Machine Learning Algorithms: Classification and Comparison.** International Journal of Computer Trends and Technology (IJCTT) –, [s. l.], v. 48, n. 3, p. 128-138, 3 jun. 2017. Disponível em: https://www.researchgate.net/profile/J-E-T-Akinsola/publication/318338750_Supervised_Machine_Learning_Algorithms_Classification_and_Comparison/links/596481dd0f7e9b819497e265/Supervised-Machine-Learning-Algorithms-Classification-and-Comparison.pdf. Acesso em: 19 jul. 2022.

OTTONI, André Luiz Carvalho; NEPOMUCENO, Erivelton Geraldo; OLIVEIRA, Marcos Santos de. **APRENDIZADO POR REFORÇO NA SOLUÇÃO DO PROBLEMA DO CAIXEIRO VIAJANTE ASSIMÉTRICO: UMA COMPARAÇÃO ENTRE OS ALGORITMOS Q-LEARNING E SARSA**. XII Simpósio de Mecânica Computacional, Diamantina, p. 87-94, 2016. Disponível em: https://www.researchgate.net/publication/309644321_Aprendizado_por_Reforco_na_Solucao_do_Problema_do_Caixeiro_Viajante_Assimetrico_Uma_comparacao_entre_os_algoritmos_Q-learning_e_SARSA. Acesso em: 4 dez. 2022.

PAIXÃO, Gabriela Miana de Mattos; SANTOS, Bruno Campos; ARAUJO, Rodrigo Martins de; RIBEIRO, Manoel Horta; MORAES, Jermana Lopes de; RIBEIRO, Antonio L. **Machine Learning na Medicina: Revisão e Aplicabilidade**. Machine Learning na Medicina, [s. l.], n. 1, ed. 118, p. 95-102, 2020. DOI https://doi.org/10.36660/abc.20200596. Disponível em: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8959062/pdf/0066-782X-abc-118-01-0095.pdf. Acesso em: 19 jul. 2022.

PARENTE, Cristina; COSTA, Daniel; SANTOS, Monica; CHAVES, Rosário Rito. **Empreendedorismo social: contributos teóricos para a sua definição**. XIV Encontro Nacional de Sociologia Industrial, das Organizações e do Trabalho Emprego e coesão social: da crise de regulação à hegemonia da globalização, Lisboa, 26 maio 2011. Disponível em: https://repositorio-aberto.up.pt/bitstream/10216/61862/2/cparenteempreendedorismo000151867.pdf. Acesso em: 14 jun. 2022.

PEREIRA, Raquel; ROLO, Ana. **O EMPREENDEDORISMO SOCIAL COMO PROMOTOR DO DESENVOLVIMENTOREGIONAL:O CASODOMICROCRÉDITO**. 14.ºWORKSHOP APDR : EMPREENDEDORISMOEDESENVOLVIMENTOREGIONAL, ESCE Instituto Politécnico de Setúbal, p. 182-197, 5 out. 2012. Disponível em:

http://apdr.pt/data/documents/atas_14_workshop_apdr_setubal.pdf#page=182. Acesso em: 27 jun. 2022.

PEREIRA, Cristina Alves dos Santos. **Lean manufacturing: aplicação do conceito a células de trabalho.** Orientador: Fernando Manuel Bigares Charrua Santos. 2010. 116 p. Dissertação (Mestrado em Engenharia e Gestão Industrial) - UNIVERSIDADE DA BEIRA INTERIOR, Covilhã, 2010. Disponível em: https://ubibliorum.ubi.pt/handle/10400.6/1921. Acesso em: 24 jul. 2022.

PIMENTEL, Edson P.; DE FRANÇA, Vilma F.; OMAR, Nizam. **A identificação de grupos de aprendizes no ensino presencial utilizando técnicas de clusterização**. SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO - SBIE, [s. l.], p. 495-504, 2003. DOI http://dx.doi.org/10.5753/cbie.sbie.2003.495-504. Disponível em: http://ojs.sector3.com.br/index.php/sbie/article/view/280. Acesso em: 25 set. 2022.

PIPE SOCIAL, **2º Mapa de Negócios de Impacto**. 2019. Disponível em: https://pipelabo.com/mapas/2o-mapa-de-negocios-de-impacto-2019/. Acesso em: 19 jun. 2022.

PIPE SOCIAL, **3º Mapa de Negócios de Impacto – Relatório Nacional**. 2021. Disponível em: https://mapa2021.pipelabo.com/downloads/3_Mapa_de_Impacto_Relatorio_Nacional.pdf. Acesso em: 19 jun. 2022.

PIRES, Ricardo; SANTOS, Sara Dereste dos; SOUZA, Miguel Angelo de Abreu de; OLIVEIRA, Lincoln Marques; MACHADO, Gabriely Augusto. **APRENDIZADO NÃO SUPERVISIONADO PARA O MAPEAMENTO DO COMPORTAMENTO TEMPORAL DE SINAIS CEREBRAIS EM EPILEPSIA**. Tekhne e Logos, Botucatu, v. 12, n. 1, p. 116-129, 2021. Disponível em: http://revista.fatecbt.edu.br/index.php/tl/article/view/765. Acesso em: 4 dez. 2022.

RAHMAN, M. Lutfor; ELAHI, Khandakar. **Micro-credit and micro-finance: functional and conceptual differences**. 20 ago. 2006. DOI: 10.1080/0961452060079248 . Disponível em: https://www.researchgate.net/publication/263170015_Micro-credit_and_micro-finance_Functional_and_conceptual_differences. Acesso em: 4 jun. 2022.

RIBEIRO, Daniel de Araújo. **Microcrédito como Meio de Desenvolvimento Social**. Orientador: Professora Drª. Miriam de Abreu Machado Campos. 2009. 105 p. Dissertação (Mestrado em Direito Empresarial) - Faculdade de Direito Milton Campos, Nova Lima, 2009. Disponível em: https://portalfat.mte.gov.br/wp-content/uploads/2016/03/DISSERTA%C3%87%C3%83O-Microcr%C3%A9dito-como-Meio-de-Desenvolvimento-Social-por-Daniel-de-Ara%C3%BAjo-Ribeiro.pdf. Acesso em: 4 dez. 2022.

SACHS, Ignacy. **Desenvolvimento: includente, sustentável, sustentado**. 1. ed. Rio de Janeiro: Garamond, 2004. 151 p.

SAKO, Mari; PARNHAM, Richard. **Technology and innovation in legal services: final report for the solicitors regulation authority**. University of Oxford, 2021. Disponível em: https://www.sra.org.uk/globalassets/documents/sra/research/fullreport-

technology-and-innovation-in-legal-services.pdf?version=4a1bfe. Acesso em: 25 set. 2022.

SARKAR, Soumajyoti; ALVARI, Hamidreza. **Mitigating bias in online microfinance platforms: A case study on Kiva. org.** Joint European Conference on Machine Learning and Knowledge Discovery in Databases, [s. l.], p. 75-91, 14 set. 2020. DOI https://doi.org/10.48550/arXiv.2006.12995. Disponível em: https://arxiv.org/abs/2006.12995. Acesso em: 16 out. 2022.

SIMÕES-GOMES, L.; ROBERTO, E.; MENDONÇA, J. Viés algorítmico – um balanço provisório. Estudos de Sociologia, [S. l.], v. 25, n. 48, 2020. DOI: 10.52780/res.13402. Disponível em: https://periodicos.fclar.unesp.br/estudos/article/view/13402. Acesso em: 16 out. 2022.

SCHAEFER, Ricardo. **Empreender como uma forma de ser, saber e fazer: o desenvolvimento da mentalidade e do comportamento empreendedores por meio da educação empreendedora.** Orientador: Prof. Dr. Ítalo Fernando Minello. 2018. 281 p. Tese (Doutorado em Administração) - Universidade Federal de Santa Maria, Santa Maria, RS, 2018. Disponível em: https://repositorio.ufsm.br/bitstream/handle/1/16023/TES_PPGADMINISTRACAO_2018_SCHAEFER_RICARDO.pdf?sequence=1&isAllowed=y. Acesso em: 19 jul. 2022.

SCHERER, Patricia Cristina. **Entendendo os negócios com impacto social : uma análise dos elementos constituintes do modelo de negócio**. Orientador: Petrini, Maira de Cássia. 2014. 114 p. Dissertação (Programa de Pós-Graduação em Administração e Negócios) - Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), [S. l.], 2014. Disponível em: http://tede2.pucrs.br/tede2/handle/tede/5688. Acesso em: 24 jun. 2022.

SEGATTO, Andréa Paula; DIÓGENES, Carla Gomes Beuter; BISCAIA, Heloisa Gappmayerr. **Tecnologias Sociais para Inclusão e Transformação Social: Caso do Microcrédito**. XV SEMEAD: Seminários em Administração, [*s. l.*], outubro 2012. Disponível em: http://sistema.semead.com.br/15semead/resultado/trabalhosPDF/363.pdf. Acesso em: 27 jun. 2022.

SHANE, Scott e VENKATARAMAN, Sankaran. **The promise of entrepreneurship as a field of research**. Academy of management review, v. 25, n. 1, p. 217-226, 2000.

SILVA, Fernando; SCHAEFFER-FILHO, Alberto E.. **Mitigando Ataques com a Orquestração de VNFs Baseadas em Contêineres Usando Aprendizado Supervisionado**. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS (SBRC), 39. , 2021, Uberlândia. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2021 . p. 350-363. ISSN 2177-9384. DOI: https://doi.org/10.5753/sbrc.2021.16732. Acesso em: 04 dez. 2022.

SIQUEIRA FILHO, Jaime Albuquerque. **Microcrédito na região metropolitana do Recife : experiência empreendedora do CEAPE**. 2005. 95 p. Dissertação (Mestrado em Economia) - Universidade Federal de Pernambuco, [S. l.], 2005. Disponível em: https://repositorio.ufpe.br/bitstream/123456789/4355/1/arquivo5967_1.pdf. Acesso em: 19 jul. 2022.

STANTON, Jeffrey; PALMER, Carole L.; BLAKE, Catherine; ALLARD, Suzie. **Interdisciplinary Data Science Education**. In: SPECIAL Issues in Data Management. ACS Symposium Series: American Chemical Society, 2012. v. 1110, cap. 6, p. 97-113. ISBN 9780841227125. Disponível em: 10.1021/bk-2012-1110.ch006. Acesso em: 17 jul. 2022.

SUN, Wenlong; NASRAOUI, Olfa; SHAFTO, Patrick. **Evolution and impact of bias in human and machine learning algorithm interaction**. PLoS ONE, [s. l.], v. 15, ed. 8, 13 ago. 2020. DOI https://doi.org/10.1371/journal.pone.0235502. Disponível em: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0235502. Acesso em: 4 nov. 2022.

TAKAKI, Patrícia; DUTRA, Moisés. **Ciência de dados na educação: contribuições interdisciplinares.** Informação, Dados e Tecnologia. Advanced Notes in Information Science, [s. l.], v. 2, p. 149-160, 2022. DOI 10.47909/ anis.978-9916-9760-3-6.94. Disponível em: https://www.researchgate.net/profile/Patricia-Takaki/publication/361152273_Data_science_in_education_interdisciplinary_contributions/links/629fd0d9c660ab61f86b4171/Data-science-in-education-interdisciplinary-contributions.pdf. Acesso em: 18 jul. 2022.

UNDP. **HUMAN DEVELOPMENT REPORT**. 2020. Disponível em: https://hdr.undp.org/en/content/human-development-report-2020?utm_source=EN&utm_medium=GSR&utm_content=US_UNDP_PaidSearch_Brand_English&utm_campaign=CENTRAL&c_src=CENTRAL&c_src2=GSR&gclid=Cj0KCQjwheyUBhD-ARIsAHJNM-N3yZBi8wgKR9M90-XsXRxcEXW6eP-fUEZYdYRFc5lU6wVDVv2IM74aAkTwEALw_wcB. Acesso em: 26 maio 2022.

VAN DYK, David; FUENTES, Montse; JORDAN, Michael I.; NEWTON, Michael; RAY, Bonnie K.; LANG, Duncan Temple. **ASA Statement on the Role of Statistics in Data Science.** Amstat News: The Membership Magazine of the American Statistical Association, [s. l.], 1 out. 2015. Disponível em: https://magazine.amstat.org/blog/2015/10/01/asa-statement-on-the-role-of-statistics-in-data-science/. Acesso em: 17 jul. 2022.

VIEIRA, Marina Lopes. **Gestão da Inovação em Negócios Sociais do Brasil.** 2019. 94 p. Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Produção) - Universidade Federal de Juiz de Fora, [S. l.], 2019.

VOMMARO, Pablo. **O mundo em tempos de pandemia: certezas, dilemas e perspectivas.** Revista Direito e Praxis, [*S. l.*], v. 12, n. 2, p. 1095-1115, 2021. DOI 10.1590/2179-8966/2020/51001| ISSN: 2179-8966. Disponível em: https://www.scielo.br/j/rdp/a/FgCBc5MG7zRNvNkWGzbgWPJ/abstract/?lang=en. Acesso em: 30 maio 2022.

WORLD INEQUALITY LAB. **World Inequality Report**. 2022. Disponível em: https://wir2022.wid.world/. Acesso em: 26 maio 2022.

YUNUS, M.; MOINGEON, B.; LEHMANN-ORTEGA, L. **Building Social Busines-Models: Lessons from the Grameen Experience**. Long Range Planning 43 (2010) 308-325. Disponível em SSN: https://ssrn.com/abstract=2673495. Acesso em: 20 jul. 2022.

ZHANG, Z.; ZHANG, D. **What is Data Science? An Operational Definition based on Text Mining of Data Science Curricula**. Journal of Behavioral Data Science, [S. l.], v. 1, n. 1, p. 1–16, 2021. DOI: 10.35566/jbds/v1n1/p1. Disponível em: https://jbds.isdsa.org/index.php/jbds/article/view/2. Acesso em: 17 jul. 2022.

ANEXO A – TERMO DE AUTENTICIDADE

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA

Termo de Declaração de Autenticidade de Autoria

Declaro, sob as penas da lei e para os devidos fins, junto à Universidade Federal de Juiz de Fora, que meu Trabalho de Conclusão de Curso do Curso de Graduação em Engenharia de Produção é original, de minha única e exclusiva autoria. E não se trata de cópia integral ou parcial de textos e trabalhos de autoria de outrem, seja em formato de papel, eletrônico, digital, áudio-visual ou qualquer outro meio.
Declaro ainda ter total conhecimento e compreensão do que é considerado plágio, não apenas a cópia integral do trabalho, mas também de parte dele, inclusive de artigos e/ou parágrafos, sem citação do autor ou de sua fonte.
Declaro, por fim, ter total conhecimento e compreensão das punições decorrentes da prática de plágio, através das sanções civis previstas na lei do direito autoral[1] e criminais previstas no Código Penal[2], além das cominações administrativas e acadêmicas que poderão resultar em reprovação no Trabalho de Conclusão de Curso.

Juiz de Fora, 23 de Junho de 2023.

Judy Inglez de Almeida — NOME LEGÍVEL DO ALUNO (A)

201749070 — Matrícula

Judy Inglez de Almeida — ASSINATURA

115.014.077-16 — CPF

[1] LEI N° 9.610, DE 19 DE FEVEREIRO DE 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências.
[2] Art. 184. Violar direitos de autor e os que lhe são conexos: Pena - detenção, de 3 (três) meses a 1 (um) ano, ou multa.

ANEXO B – SOLICITAÇÃO DE ACESSO AO CÓDIGO KIVA

8 de 13

Your comment:
Dear Mr./Ms.,

My name is Judy, I'm graduating in Production Engineering at UFJF (Minas Gerais, Brazil).

I am contacting you because, at the moment, I am writing my course conclusion work, and I would like to use the Kiva website algorithm as an object of study.

My goal is to analyze and classify the algorithm used by Kiva.org within the four types of Machine Learning analysis (supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning), highlighting the importance of data science for business/social enterprises.

In addition, I aim to evaluate how the algorithmic bias can interfere with the options that are shown to the final consumer, and show some options for resolutions already found in the academic environment.

My work seeks to unite three fields of Engineering: social entrepreneurship, innovation through data science and social responsibility. Given this, I believe that the Kiva site is a great object of work.

Therefore, I would like to know if there is any possibility of having access to the code used by you, only for academic reasons, always referencing and giving due credit. Could you share the code used on Kiva.org?

Thank you very much in advance,
Yours sincerely,
Judy Thuler de Almeida

8 de 13

Re: Kiva as an object of study [ref:_00D506svl._5001T1jxGBl:ref] Externa 2023.1/TCC ×

Kiva Community Support <contactus@kiva.org> seg., 19 de set. de 2022, 20:58
para mim

inglês > português Traduzir mensagem Desativar para: inglês ×

Hi Judy,

Thank you so much for reaching out. We're so happy to hear that you're interested in learning more about Kiva for your school project.

Unfortunately, Kiva is usually unable to fulfill specific research requests. However, we encourage you to fill out our research request form, and someone may get back to you if your request is possible: https://docs.google.com/forms/d/1n7sUZDgQTLaDsDVzUOuFNm-kT8bjo3kB93fLvq5QqnY/viewform?edit_requested=true

Additionally, here are a few pages of our site that may provide useful information for your research:
--kiva.org/about
--cgap.org
--mftransparency.org
--mixmarket.org

Many researchers have also found our API to be a valuable resource. You can learn more about what information is available via our API at: kiva.org/build

Lastly, students and educators have also found the Kiva U resources and materials helpful as they look to incorporate Kiva into their academic programs and studies. For more information, check out: KivaU.org

Thank you for your interest in including Kiva in your study, and it would be great to hear what results you eventually find through your research. Please feel free to send us a link to your published article.

If I can help with anything else please let me know.

Best wishes,
Franklin Lester
Community Support Agent

Kiva.org Loans That Change Lives