

Previsão do preço do carbono por modelos de aprendizado de máquina

Carbon price prediction by machine learning models

Jorge Yoshio Kanda ¹

André Carlos Ponce de Leon Ferreira de Carvalho ²

Resumo

O controle e a redução das emissões de gases de efeito estufa são ações necessárias para evitar as possíveis consequências provenientes das mudanças climáticas. Neste contexto, o mercado de carbono tem grande relevância, principalmente, para países em desenvolvimento como o Brasil que possui uma imensa riqueza ambiental com a sua Floresta Amazônica. O propósito deste estudo é identificar conjuntos de dados de sustentabilidade a serem usados por modelos preditivos de aprendizado de máquina (AM) capazes de estimar com a maior precisão possível o preço do carbono praticado no mercado mundial. Em nossos experimentos computacionais, algoritmos foram implementados a partir de diferentes algoritmos de AM, sendo usados como parâmetros de entrada diversos conjuntos de dados. Os resultados obtidos mostram que dados amazônicos parecem ter uma relação direta com o preço do carbono praticado no mercado mundial. Um procedimento de seleção de atributos foi aplicado na união dos conjuntos dos dados amazônicos que também foi submetido aos mesmos modelos de AM para verificar se há melhorias no desempenho preditivo. Portanto, havendo uma estimativa precisa do preço do carbono e o Brasil regulamentando as regras para o comércio do carbono, a Região Amazônica tende a ser beneficiada com ganhos significativos nos aspectos ambientais, econômicos e sociais.

Palavras-chaves: Crédito de Carbono; Inteligência Artificial; Modelos de Regressão; Sustentabilidade.

Abstract

Controlling and reducing greenhouse gas emissions are necessary actions to avoid the possible consequences of climate change. In this context, the carbon market is of great relevant, especially for developing countries, such like Brazil, which has an immense environmental wealth with its Amazon Forest. The purpose of this study is to identify sustainability datasets to be used by predictive machine learning (ML) models that are able to accurately estimate the price of carbon practiced in the main world market. In our computational experiments, algorithms were implemented from different ML algorithms, using different datasets as input parameters. The results obtained show that Amazonian data seem to have a direct relationship with the price of carbon practiced in the world market. A feature selection procedure was applied to the union of the Amazonian datasets and submitted to the same ML models to verify if there are improvements in the predictive performance. Therefore, with an accurate estimate of the carbon price and Brazil regulating the rules for carbon trading, the Amazon Region tends to benefit from significant gains in environmental, economic, and social aspects.

Keywords: Carbon Credit; Artificial Intelligence; Regression Models; Sustainability.

Recebido em (*manuscript first received*): 15/03/2023

Aprovado em (*manuscript accepted*): 15/05/2023



DOI: <http://dx.doi.org/10.17648/aos.v12i2.2916>

1 Introdução

A pandemia da Covid-19 provocou fortes impactos econômicos e sociais no mundo inteiro, sendo que no Brasil houve o agravamento da crise econômica em um quadro social que já era preocupante (Costa & Soares, 2020). O planejamento da recuperação da economia brasileira deveria conter planos bem delineados que possam promover o desenvolvimento sustentável. Para isso, é fundamental que sejam estabelecidos mecanismos de regulação e instrumentos de precificação do carbono no país (Eibel & Pinheiro, 2015).

¹ Doutorado em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo, USP, Brasil. Professor Associado pela Universidade Federal do Amazonas, UFAM, Brasil E-mail: jkanda@ufam.edu.br

² Pós-doutorado pela University of Kent of Canterbury, UKC, Grã-Bretanha. Pós-doutorado pela Universidade do Porto, Porto, Portugal. Livre-docência pela Universidade de São Paulo, USP, Brasil. Professor pela Universidade de São Paulo, USP, Brasil. E-mail: andre@icmc.usp.br

Nas últimas décadas, o planeta Terra tem passado por mudanças climáticas inesperadas, que geram apreensões de como será o mundo para as gerações futuras. Em função disso, muitos governos, organizações, empresas e indivíduos têm se mobilizado para tentar conter principalmente as emissões de gases de efeito estufa (GEEs) de origem antropogênica (Piva *et al.*, 2021). De acordo com Fan e Todorova (2017), existem países que já começaram a adotar algum tipo de mecanismo para a comercialização do carbono com a finalidade de controlar as emissões de dióxido de carbono (CO₂), metano (CH₄) e óxido nitroso (N₂O).

Na visão de Kim e Huh (2020), o comércio de créditos de carbono deve proporcionar uma redução da emissão de GEEs em diversos países, que por consequência haverá menos poluição no meio ambiente em tais lugares. Isso deve acontecer principalmente a partir de incentivos à criação e desenvolvimento de projetos tecnológicos para os quais tenham sido previstas execuções de atividades econômicas mais limpas e, assim sendo, haverá uma contribuição significativa para uma maior proteção ambiental.

Empresas e países em desenvolvimento que investem em projetos de mecanismo de desenvolvimento limpo (MDL) podem obter créditos de carbono, sendo que cada crédito corresponde a uma tonelada de dióxido de carbono não emitida na atmosfera (Eibel & Pinheiro, 2015). Os detentores dos créditos podem negociá-los no mercado de carbono junto às empresas e aos países desenvolvidos que não conseguem alcançar as suas metas assumidas para reduzir a emissão dos GEEs (Oliveira, 2018). O mercado de carbono da União Europeia é o mais antigo contendo registro de cotação do crédito de carbono desde o dia 25 de março de 2008 (ICAP, 2022).

No Brasil, as principais oportunidades de geração de créditos de carbono estão nos setores de florestas, energia e tratamento de resíduos. Em tais setores, as possibilidades de mitigações de emissões de GEE podem proporcionar novas oportunidades econômicas para o país com a estimativa de serem gerados entre US\$ 1,76 bilhões e US\$ 5,87 bilhões no ano de 2030 a partir da oferta de carbono no mercado mundial (Pinheiro *et al.*, 2022).

Conforme a análise feita pelo Cebds (2021), o processo de reflorestamento a partir do manejo e restauração florestal é uma das melhores formas para mitigar os efeitos das emissões no país. Todavia, percebe-se que as ações de preservação do meio ambiente, por meio da redução do desmatamento e da emissão de GEE possuem um efeito mais imediato. Nesse contexto, preservar a Floresta Amazônica é muito importante, pois ela tem a capacidade de armazenar bilhões de toneladas de carbono em sua biomassa, além de abrigar inúmeras espécies de árvores e plantas, fornecendo uma série de serviços ecossistêmicos de inestimável valor local, regional e global (Galbraith *et al.*, 2014).

Salienta-se que a preocupação com as alterações do meio ambiente decorrentes do aumento da quantidade de dióxido de carbono na atmosfera não é recente. Fearsinde (1985) já afirmava que a devastação contínua da Floresta Amazônica e o uso de combustível fóssil estavam contribuindo para o aquecimento da Terra entre 1,5°C e 4,5°C antes do ano de 2050. Na época, ele reconheceu que esta estimativa pode não ter sido tão precisa devido à escassez de dados confiáveis a respeito da taxa de devastação da Amazônia Brasileira e da quantidade de carbono liberado na região.

Atualmente, “os dados brutos de inventários florestais no Brasil ainda são escassos ou encontram-se de maneira dispersa em diversas bases de dados, muitas vezes com acesso restrito, limitando a realização de estudos interdisciplinares” (Piva *et al.*, 2021, p. 3). De fato, dados históricos relacionados às mudanças climáticas podem ser obtidos a partir de diferentes fontes, tais como: o Sistema de Registro Nacional e Emissões (Brasil, 2017); o Sistema de Estimativa de Emissões de Gases de Efeito Estufa por Município (SEEG, 2022); os indicadores dos municípios amazonenses em relação a cada um dos 17 Objetivos de Desenvolvimento Sustentável definidos pela Organização das Nações Unidas (UFAM, 2019); e o Sistema de Monitoramento de Queimadas (INPE, 2020).

Quando os dados são produzidos continuamente em grande volume, originados de fontes diversas e dispostos em diferentes formatos, tem-se o que se chama de *Big Data* (Chen *et al.*, 2014). Produzir uma análise das informações contidas em um *Big Data* de modo que possam ser úteis para auxiliar um gestor em suas tomadas de decisões é bastante desafiador (McAfee & Brynjolfsson, 2012). Por isso, nessas condições, as decisões tendem a serem mais efetivas se houver auxílio de

recursos computacionais para armazenar, processar, analisar e encontrar possíveis relações entre os dados disponíveis. A identificação de padrões que possam existir em um conjunto de dados pode ser feita de forma mais rápida e acurada por modelos preditivos baseados em inteligência artificial (IA) que sejam induzidos a partir de técnicas de aprendizado de máquina (AM) (Tan *et al.*, 2006).

Diante desse contexto, o objetivo principal deste trabalho consiste em identificar conjuntos de dados de sustentabilidade a serem usados por modelos preditivos induzidos por algoritmos de AM que sejam capazes de estimar com a maior precisão possível o preço do carbono praticado em um dos principais mercados do mundo.

O restante deste artigo está estruturado conforme segue. A Seção 2 apresenta os conceitos dos principais termos relacionados à pesquisa, bem como uma breve explanação dos principais trabalhos relacionados ao tema explorado. Na Seção 3, estão detalhados os aspectos metodológicos adotados durante o desenvolvimento das atividades deste estudo. As descrições dos experimentos computacionais; assim como, os resultados experimentais e as análises realizadas estão dispostas na Seção 4. E, finalmente, as considerações finais seguem na Seção 5, com as conclusões e sugestões de trabalhos futuros para a continuidade da pesquisa.

2 Fundamentação Teórica

A preocupação em se ter um mundo mais ecológico tem impulsionado o aumento de pesquisas em computação aplicada envolvendo outras áreas do conhecimento como, por exemplo, o meio ambiente. Uma proposta para o uso das expressões: “Inteligência Artificial Sustentável”, “Inteligência Artificial para a sustentabilidade” e “Sustentabilidade da Inteligência Artificial” foram apresentadas por (Wynsberghe, 2021). O termo sustentabilidade deve ser entendido pelo conjunto de ações que envolvem os aspectos ambiental, social e econômico (Mikhailova, 2004).

Zhang e Huisingh (2017) recomendam o uso de mineração de dados e *big data* para selecionar locais geológicos de CO₂, estimar as capacidades de armazenamento de tais locais, avaliar os potenciais riscos de vazamento e analisar a concessão de créditos de carbono. Zhou *et al.* (2019) afirmam que é difícil prever antecipadamente com elevado índice de acerto o preço do carbono a ser praticado no mercado devido às suas características não lineares e não estacionárias.

De acordo com Zhang *et al.* (2018), existem duas teorias principais para fazer a previsão do preço do carbono: uma é produzir uma previsão de fatores baseada em elementos como: carvão, petróleo, gás natural, preços de energia e índices de ações globais, porém ressaltam que esta forma pode resultar em falhas na previsão do preço do carbono. A outra teoria consiste em realizar a previsão fazendo uso de séries temporais para a construção de modelos estatísticos de previsão econométrica e modelos baseados em IA, como de modelos gerados por algoritmos de AM, que se baseiam em dados históricos de preço para produzir uma previsão futura.

Kaack *et al.* (2020) salientam que atualmente existe muito interesse por parte dos pesquisadores em tentar entender os efeitos produzidos por um processo de AM na ação climática. Eles acreditam que essa busca por projeções do clima pode ajudar a ter políticas públicas mais adequadas para a sociedade. Kaack *et al.* (2022) afirmam que os impactos mais fáceis de medir não são aqueles com os maiores efeitos, sendo que o maior desafio está em estimar os efeitos de macroescala. Com isso, sugerem que se deva partir das tendências subjacentes a esses efeitos, priorizando as ações que alinhem os estudos de AM às estratégias climáticas. Kanda e Carvalho (2022) realizaram um estudo preliminar do uso de IA para prever o preço do carbono.

Nas subseções que seguem são explanados os principais conceitos relacionados ao tema da pesquisa em desenvolvimento.

2.1 Gases de Efeito Estufa

De acordo com o Sistema de Estimativa de Emissões de Gases de Efeito Estufa - SEEG (2019), o escopo dos gases de efeito estufa incluem todos os gases previstos nos inventários dentre os quais estão o

dióxido de carbono (CO₂), metano (CH₄) e óxido nitroso (N₂O), que juntos respondem por mais de 99% das emissões em carbono equivalente (CO₂e).

Existem duas abordagens principais para determinação do carbono equivalente: *Global Warming Potential* (GWP) e *Global Temperature Change Potential* (GTP). A primeira considera a influência dos gases na alteração do balanço energético da Terra e a segunda avalia a influência no aumento de temperatura. Ambas são medidas para um prazo de 100 anos, todavia o mais comumente utilizado é GWP (SEEG, 2019). GWP mensura quantas vezes mais calor uma determinada quantidade de um GEE retém na atmosfera em relação a uma mesma quantidade de CO₂ em um dado intervalo de tempo. Essa medida é expressa como um fator que, multiplicado pela massa do gás, resulta em uma massa equivalente de CO₂, denominado de gás carbônico equivalente (CO₂e) (Brasil, 2021).

Tabela 1 - Transformação de Gases de Efeito Estufa em Gás Carbônico Equivalente

GÁS	SÍMBOLO	GTP-100	GWP-100
DIÓXIDO DE CARBONO	CO ₂	1	1
METANO	CH ₄	4	28
ÓXIDO NITROSO	N ₂ O	234	265
HIDROFLUORCARBONETOS	HFC- 23	12.700	12.400
	HFC - 125	967	3.170
	HFC - 134a	201	1.300
	HFC - 143a	2.500	4.800
	HFC - 152	19	138
PERFLUORCARBONETOS	CF ₄	8.040	6.630
	C ₂ F ₆	13.500	11.100
HEXAFLUORETO DE ENXOFRE	SF ₆	28.200	23.500

Fonte: Assis & D'Agosto (2020).

Para equiparar os efeitos produzidos pelos diferentes GEEs devem ser utilizadas métricas que transformam o valor de cada um dos gases em CO₂e, multiplicando-o pelo fator de equivalência mostrado na Tabela 1 (Assis & D'Agosto, 2020).

2.2 Crédito de Carbono

Créditos de carbono são certificados emitidos aos projetos que comprovadamente reduzem a emissão de gases que causam o efeito estufa (Oliveira, 2018, p. 95). Nesse sentido, empresas e países em desenvolvimento que investem em projetos de mecanismo de desenvolvimento limpo (MDL) podem obter créditos de carbono e, conseqüentemente, negociá-los no mercado com empresas e países desenvolvidos que estão poluindo acima de um limite pré-definido, não alcançando as suas metas obrigatórias de redução de emissão dos GEEs (Eibel & Pinheiro, 2015).

O valor do limite foi inicialmente definido em 1997 no Protocolo de Kyoto, onde ficou acordado que os países do Anexo I (países desenvolvidos) deveriam reduzir suas emissões de GEE no período entre 1998 e 2012 em pelo menos 5,2% em relação ao nível de 1990 (Dias *et al.*, 2020). Recentemente, a Conferência das Nações Unidas sobre as mudanças climáticas de 2021 - COP 26 decidiu que deverá haver uma redução global das emissões de dióxido de carbono (CO₂) em 45% até o ano de 2030, comparado aos valores de 2010. Além disso, na mesma conferência também ficou decidido pela neutralidade da liberação do CO₂ até 2050, quando as emissões devem ser reduzidas ao máximo e as restantes totalmente compensadas por reflorestamento e tecnologias de captura de carbono da atmosfera (Pinheiro *et al.*, 2022). Ademais, os países participantes nesse evento assumiram o compromisso de zerar o desmatamento e reduzir a emissão de metano em 30% até 2030 (Passarinho, 2021).

2.3 Mercado de Carbono

O mercado comercial de carbono é uma ferramenta importante para frear as mudanças climáticas globais (Dormady, 2014). Em decorrência do grande volume de emissões de GEEs que ainda ocorrem no mundo, o mercado de carbono é um local relevante onde o comportamento humano em relação ao meio ambiente é formulado como problema a ser resolvido na sociedade atual. É um lugar de abrangência global que foca nas atividades ecológicas economicamente diferenciadas. Esse mercado não pode funcionar sem quantificar as emissões de GEEs (Whittington, 2016).

Do ponto de vista de Whittington (2016), a compensação do carbono é um tipo de produto financeiro, mas determinar o seu valor pode ser considerado um negócio complicado. Considerando esse aspecto, para que ocorra uma efetiva negociação dos créditos de carbono, é necessário estabelecer políticas, estratégias e tecnologias capazes de mensurar a comercialização dos créditos de carbono (Kim & Huh, 2020). Na projeção feita por Kim e Huh (2020), o mercado de créditos de carbono se tornará uma questão crítica a partir de 2021, de modo que os créditos de carbono serão aplicados a sistemas onde os indivíduos possam negociar entre si.

Segundo Albuquerque *et al.* (2021), o volume emitido de créditos de carbono no mercado mundial em 2019 foi de 142 MtCO₂e (milhões de toneladas de dióxido de carbono equivalente), sendo que 3,6% desse valor foram emitidos no Brasil. Conforme dados do Conselho Empresarial Brasileiro para o Desenvolvimento Sustentável, entre os anos de 2019 e 2020, os créditos por ativos florestais cresceram no Brasil, de 3 MtCO₂ para 15 MtCO₂ (Cebds, 2021). A oferta de créditos de carbono tende a ser maior nos setores de Florestas e Agropecuária por causa das medidas de mitigação de emissões terem melhores relações de custo-efetividade. Isso sem considerar que o setor de Energia tem grande potencial de mitigação para o país em razão da inovação tecnológica e enorme experiência em projetos de MDL (Cebds, 2021).

Durante o Protocolo de Kyoto foram definidos dois mecanismos de compensação para o mercado de crédito de carbono: a implementação conjunta e o MDL. A implementação conjunta é caracterizada quando a negociação é realizada entre duas ou mais empresas ou países do Anexo I. Já o MDL só ocorre se o comércio for entre uma empresa ou país desenvolvido e uma empresa ou país em desenvolvimento (Albuquerque *et al.*, 2021).

A respeito do MDL, países em desenvolvimento fornecem créditos de carbono aos países do Anexo I e estes, por sua vez, alcançam suas metas de forma menos onerosa fazendo investimentos em projetos de redução de GEE, de tal maneira que o desenvolvimento sustentável é promovido nos países que recebem os projetos (Albuquerque *et al.*, 2021). Esses projetos devem ser dotados de tecnologias que reduza a poluição do meio ambiente, mitigando as mudanças climáticas. Assim sendo, a economia local tende a crescer com a geração de mais empregos e, conseqüentemente, deve haver uma melhoria considerável nas condições de vida da população local. De acordo com a UNFCCC (2018), o MDL tem contribuído com cerca 32 bilhões de dólares em investimentos relacionados a 340 projetos desenvolvidos no Brasil.

O Governo Federal Brasileiro, por meio do Decreto nº. 11.075, de 19 de maio de 2022, estabeleceu procedimentos para a elaboração dos planos setoriais de mitigação das mudanças climáticas com a instituição do Sistema Nacional de Redução de Emissões de Gases de Efeito Estufa - SINARE. Esse sistema tem a finalidade de servir como central única de registro de emissões, remoções, reduções e compensações de gases de efeito estufa e de atos de comércio, de transferências, de transações e de aposentadoria de créditos certificados de redução de emissões (Brasil, 2022). Infelizmente, este decreto ainda não regula de forma efetiva a criação de um mercado de carbono regulado no Brasil.

2.4 Precificação do Carbono

A precificação do carbono é uma poderosa ferramenta política e econômica a ser usada por governos na definição de uma estratégia climática mais ampla a fim de incentivar mudanças nos padrões de produção, consumo e investimento, sustentando desse modo o crescimento de baixo carbono. Os instrumentos de precificação de carbono podem ser agrupados em duas categorias: precificação indireta e direta de carbono (WorldBank, 2022).

A precificação indireta do carbono ocorre quando há mudanças do preço dos produtos não diretamente proporcional às emissões de carbono. Por exemplo, os impostos especiais sobre o consumo de combustível que aplicam um valor de imposto fixo à gasolina por litro influenciam indiretamente no preço das emissões de carbono da combustão dessa gasolina. De maneira inversa, os subsídios que reduzem o preço dos combustíveis fósseis criam um sinal indireto “negativo” do preço do carbono, incentivando um maior consumo e, conseqüentemente, aumentando as emissões de carbono. Todos os instrumentos de política que focam no incentivo de preços para o uso de combustíveis podem ser considerados preços indiretos de carbono (WorldBank, 2022).

Sobre a precificação direta, há três tipos de instrumentos: tributo sobre o carbono, sistema de comércio de emissões e mecanismo de crédito de carbono (WorldBank, 2022), que estão detalhados a seguir:

- Tributo sobre o carbono (*carbon tax*): é um instrumento de política de governo que define uma taxa sobre as emissões de GEE (normalmente sobre uma tonelada métrica de dióxido de carbono equivalente). Neste caso, o mercado é quem determina o nível de redução de emissões de GEE incentivado pelo preço;
- Sistema de comércio de emissões (*Emission Trading System - ETS*): primeiramente é definido um valor máximo no volume total de emissões de GEE em um ou mais setores da economia. A partir disso, o governo leiloa ou distribui licenças de emissão negociáveis para entidades cobertas pelo limite, em que cada licença representa o direito de emitir um certo volume de gases. As entidades podem optar por comprar licenças adicionais, se necessário, ou vender licenças excedentes. Esse tipo de política é conhecido como sistema “*cap-and-trade*”. Nesse instrumento, o preço do carbono não é fixado pelo governo, mas determinado pelo mercado por meio da oferta e demanda de licenças ou créditos de emissão.
- Mecanismo de crédito de carbono: trata-se de um sistema em que os créditos negociáveis são gerados por meio de atividades de redução ou remoção de emissões implementadas voluntariamente. Tipicamente cada crédito representa uma tonelada métrica de dióxido de carbono equivalente. Os mecanismos de crédito de carbono operam de forma diferente dos impostos de carbono e ETS. Em vez de exigir que as empresas paguem pela emissão (ou seja, o princípio do poluidor-pagador), as empresas e outras organizações podem gerar créditos de carbono (e, portanto, receita) demonstrando que as emissões foram reduzidas em relação a uma linha de base referenciada.

2.5 Aprendizado de Máquina

Um processo de AM pode ser baseado em um paradigma de aprendizado supervisionado ou de aprendizado não-supervisionado. No primeiro, um conjunto de instâncias descritas por valores de seus atributos e do valor da sua classe é usado para treinar um modelo preditivo que seja capaz de identificar uma função entre os valores dos atributos preditivos (valores de entrada) e o valor de um atributo alvo, no caso a da classe (valor de saída). O modelo induzido pode ser utilizado para prever o valor da classe para exemplos não visto durante o seu treinamento. Em um aprendizado não-supervisionado, não se tem o valor da classe para as instâncias, por isso, é induzido um modelo descritivo, que pode, por exemplo, calcular um valor de similaridade entre as instâncias do conjunto com a finalidade de criar grupos de exemplos semelhantes entre si. Neste caso, uma nova instância será associada ao grupo que contenha exemplo(s) mais similar(es) com essa nova instância (Tan *et al.*, 2006).

A capacidade de generalização de um modelo de AM preditivo pode ser afetada pela existência de instâncias com valor ausente, incorreto, inconsistente ou duplicado. Além disso, a quantidade insuficiente de exemplos ou de atributos e a interdependência entre atributos também podem prejudicar a capacidade de aprendizado do modelo. Para maximizar a eficácia do modelo treinado, é recomendável que seja feito um pré-processamento no conjunto de dados com a finalidade de melhorar a qualidade dos dados antes de serem usados no processo de indução do modelo de AM (Faceli *et al.*, 2011).

O desempenho preditivo de um modelo de AM pode ser mensurado por meio de medidas apropriadas como a acurácia ou a taxa de erros quando se tem modelos de classificação. Estas medidas comparam o valor predito pelo modelo para cada uma das instâncias com o seu respectivo valor verdadeiro (Tan *et al.*, 2006). Se o modelo preditivo é de regressão, então o seu desempenho pode ser mensurado por meio de medidas como Erro Quadrático Médio e Distância Absoluta Média (Faceli *et al.*, 2011).

2.5.1 Seleção de meta-atributos

Quando um conjunto de dados possui muitos atributos pode ser que alguns deles sejam irrelevantes ou redundantes, e por consequência poderá influenciar negativamente o desempenho do modelo ao prever o valor de saída para novos exemplos. Segundo Mitchell (1997), atributos irrelevantes são aqueles que não possuem informações úteis que contribuem para o aprendizado do modelo. Para exemplificar, o uso do atributo "Nome do Estado" no conjunto de dados de municípios amazonenses não será útil porque todas as instâncias teriam o mesmo conteúdo. Atributos redundantes são aqueles que não necessariamente possuem valores iguais para cada instância, mas o valor de um atributo pode ser calculado a partir do valor do outro atributo (Tan *et al.*, 2006). Por exemplo, o atributo "Area Nao Desmatada" é redundante se os valores para todas as instâncias do conjunto de dados forem iguais ao resultado da subtração entre os valores dos atributos "Area Territorial" e "Area Desmatada".

Faceli *et al.* (2011) indicam que os atributos mais relevantes podem ser selecionados a partir da ordenação da relevância dos atributos ou da seleção de um subconjunto de atributos considerados relevantes. Seguem mais detalhes de cada procedimento de seleção.

No processo de ordenação, os atributos são ordenados por relevância medida conforme algum critério pré-estabelecido. Em tarefas de regressão, um atributo é indicado relevante quando pode prever, com a maior precisão possível, o valor da saída desejada. O aspecto positivo da ordenação está no tempo computacional linear em relação ao número de atributos. Entretanto, alguns atributos podem ser menos relevantes quando usados isoladamente, e mais relevantes quando selecionados em composição com outros (Faceli *et al.*, 2011).

A vantagem de selecionar um subconjunto de atributos é que a correlação entre os atributos do subconjunto não é desprezada. Por outro lado, o custo computacional cresce com o aumento da quantidade de atributos, podendo tornar intratável quando esse número é muito grande. Outro aspecto importante a ser destacado é que um subconjunto selecionado com m atributos dificilmente será constituído pelos m melhores atributos indicados pela seleção de ordenação (Tan *et al.*, 2006).

3 Materiais e Métodos

Toda e qualquer pesquisa científica deve ser realizada para buscar uma solução para um dado problema. Por isso, é necessário que sejam usados procedimentos científicos com a finalidade de descobrir fatos que sejam inerentes a uma determinada realidade (Silveira & Córdova, 2009). A seguir serão apresentados os aspectos metodológicos da pesquisa realizada quanto à abordagem, natureza, objetivos e procedimentos.

Sobre a abordagem, a pesquisa é qualitativa, já que buscou-se produzir informações sobre créditos de carbono na Região Amazônica a partir de dados ambientais, sociais e econômicos. Para isso está sendo identificada a relevância dos atributos contidos em conjuntos de dados de sustentabilidade usados por modelos computacionais baseados em AM capazes de prever o preço do carbono. Minayo (2014) argumenta que a pesquisa qualitativa trabalha com o universo de significados, motivos, aspirações, crenças, valores e atitudes, que estão em um espaço mais profundo das relações, dos processos e fenômenos que não podem ser reduzidos à mera operacionalização de variáveis.

Em relação à natureza, e em conformidade ao que estabelecem Silveira e Córdova (2009), a pesquisa é considerada básica porque estão sendo produzidos novos conhecimentos que podem contribuir para o avanço da Ciência, envolvendo verdades e interesses universais. Os conhecimentos produzidos poderão ser usados por autoridades governamentais para auxiliar em suas tomadas de decisão. E nesse

contexto, pode-se afirmar que a pesquisa é também aplicada já que abrange estudos elaborados com a finalidade de resolver problemas no âmbito da sociedade em que o pesquisador vive (Gil, 2010).

No que tange aos objetivos da pesquisa, possui uma abordagem explicativa, uma vez que estão sendo buscadas possíveis explicações nas relações existentes entre os diferentes dados coletados sobre sustentabilidade e preço de carbono, que de alguma forma poderão facilitar a adoção de medidas preventivas para reduzir os impactos negativos à natureza. Segundo Gil (2010), em uma pesquisa explicativa devem ser identificados fatores que determinam ou contribuem para a ocorrência de fenômenos.

Os procedimentos da pesquisa podem ser caracterizados como documental e experimental. Em uma pesquisa documental são consultadas diversas fontes que geralmente estão dispersas, sem a realização de um tratamento analítico (Gerhardt, 2009). A procura por conceitos, dados e informações relacionados ao tema desta pesquisa ocorreu a partir de leituras em publicações científicas e documentos oficiais de órgãos públicos e não governamentais. A pesquisa é também experimental porque foram desenvolvidos experimentos computacionais, com a implementação de modelos de regressão baseados em AM capazes de prever o preço do carbono.

Uma das primeiras providências para realizar um processo de AM é obter um conjunto de dados representativo do problema investigado, que será usado para o treinamento e teste do modelo preditivo. Nesta pesquisa, foi utilizado o conjunto de preços de carbono cotados no mercado europeu no período de 25 de março de 2008 a 30 de junho de 2022 extraído da página web do ICAP (2022).

Para os modelos preditivos estimarem o preço do carbono foram usados diferentes conjuntos de dados ambientais, econômicos e sociais como parâmetros de entrada dos modelos. Em todos os conjuntos foram realizados procedimentos de pré-processamento que resultaram nos conjuntos de metadados. As quantidades de instâncias e de meta-atributos de cada conjunto de metadados estão dispostas na Tabela 2.

Tabela 2 - Quantidade de Instâncias e Meta-atributos dos Conjuntos de Metadados

ID	METADADOS	QUANTIDADE DE INSTÂNCIAS	QUANTIDADE DE META-ATRIBUTOS
D1	GEE Mundo	2619	8
D2	GEE Brasil	1319	7
D3	GEE Amazonas	744	3
D4	Fatores de Emissão CO2 Brasil	5180	4
D5	Desmatamento Bioma Amazônia	1770	5
D6	Desmatamento Estado Amazonas	868	12
D7	Queimadas Brasil	4644	4
D8	Indicadores Sociais Brasil	243	60
D9	PIB Municípios Amazonenses	744	14

Fonte: Elaborada pelos autores (2023).

A periodicidade e o intervalo de tempo das instâncias de cada conjunto de metadados podem ser conferidos na Tabela 3. Todos os conjuntos de metadados estão publicamente disponíveis em: <https://github.com/Kanda2023?tab=repositories>.

Tabela 3 - Periodicidade dos Conjuntos de Metadados

ID	METADADOS	FREQUÊNCIA	INÍCIO	FIM
D1	GEE Mundo	Anual	2008	2018
D2	GEE Brasil	Anual	2008	2016
D3	GEE Amazonas	Anual	2008	2019
D4	Fatores de Emissão CO2 Brasil	Diário	26/03/2008	30/06/2022
D5	Desmatamento Bioma Amazônia	Diário	01/04/2008	15/11/2021
D6	Desmatamento Estado Amazonas	Anual	2008	2021
D7	Queimadas Brasil	Mensal	Mar. 2008	Jun. 2022
D8	Indicadores Sociais Brasil	Anual	2012	2020
D9	PIB Municípios Amazonenses	Anual	2008	2019

Fonte: Elaborada pelos autores (2023).

4 Experimentos Computacionais

Os modelos de regressão baseados em AM foram implementados a partir das técnicas de redes neurais artificiais, k-vizinhos-mais-próximos, árvore de decisão e floresta aleatória, que estão dentre as mais conhecidas pela comunidade científica (Faceli *et al.*, 2011). Além destes, também foram criados um modelo de regressão linear e um modelo *baseline* para efeitos de comparação de desempenho entre os diferentes modelos.

A implementação dos modelos foi feita usando o *software Python* versão 3.10.6 (Rossum & Drake, 2009) e a ferramenta *scikit-learn* versão 1.1.2 (Pedregosa *et al.*, 2011). Dessa ferramenta foram escolhidos os pacotes *LinearRegression*, *MLPRegressor*, *KNeighborsRegressor*, *DecisionTreeRegressor* e *DummyRegressor* para induzir os modelos de regressão que denominamos de: LR, MLP, KNN, DT e *baseline*, respectivamente. A estratégia do modelo *baseline* é bem simples, o valor predito é sempre a média dos valores do conjunto de treinamento (Pedregosa *et al.*, 2011). O pacote *RandomForestRegressor* foi também utilizado posteriormente na implementação de um modelo preditivo que combina a saída de vários modelos preditivos, denominado comitê ou *ensemble*.

Durante o treinamento dos modelos preditivos foi adotado o procedimento de validação cruzada em 10 pastas, que consiste em distribuir igualmente as instâncias em 10 pastas. Em cada execução do modelo, uma pasta diferente é usada para validar a capacidade preditiva do modelo após ter sido treinado com as instâncias das demais pastas (Tan *et al.*, 2006). O desempenho preditivo dos modelos nos conjuntos de treinamento foi mensurado por meio da medida Erro Quadrático Médio, para a qual quanto menor o valor, melhor é o desempenho.

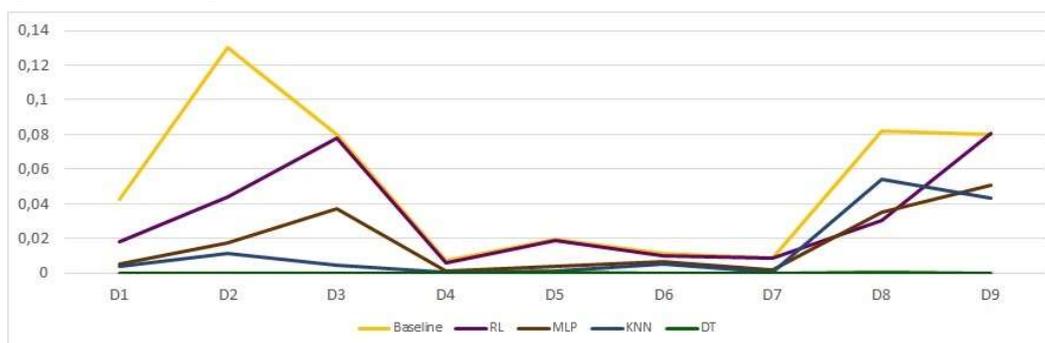
Todos os algoritmos foram executados em um *notebook* contendo um processador Intel(R) Core (TM) i5-7200U CPU@2.50GHz 2.71 GHz, 4 GB de memória RAM e Sistema Operacional *Windows 10 Single Language* Versão 21H1 de 64 bits.

Com a finalidade de simular um cenário para os modelos treinados preverem o preço do carbono para novos exemplos, as instâncias mais recentes dos últimos 12 meses de cada conjunto de metadados foram separadas para testar a capacidade preditiva dos modelos. Após o treinamento, o conjunto de teste de cada metadados foi submetido aos modelos, que forneceram um valor do preço de carbono para cada instância nova. Este valor foi comparado com o valor verdadeiro e a Distância Absoluta Média foi mensurada para cada modelo após processar todas as instâncias do conjunto de teste.

4.1 Resultados Experimentais

A Figura 1 mostra o desempenho dos modelos mensurado pelo erro quadrático médio em cada conjunto de metadados utilizado durante o processo de treinamento. Percebe-se que todos os modelos de AM tiveram melhor desempenho do que o modelo *baseline* (linha amarela), exceto o modelo baseado em regressão linear (LR) que para determinados conjuntos de metadados o desempenho foi igual à do *baseline*. A inferioridade da capacidade preditiva do modelo LR decorre, muito provavelmente, de que a função que melhor mapeia os valores de entrada para o valor de saída, considerando os conjuntos de metadados investigados, é não-linear.

Figura 1 - Desempenho dos Modelos Preditivos nos Conjuntos de Treinamento



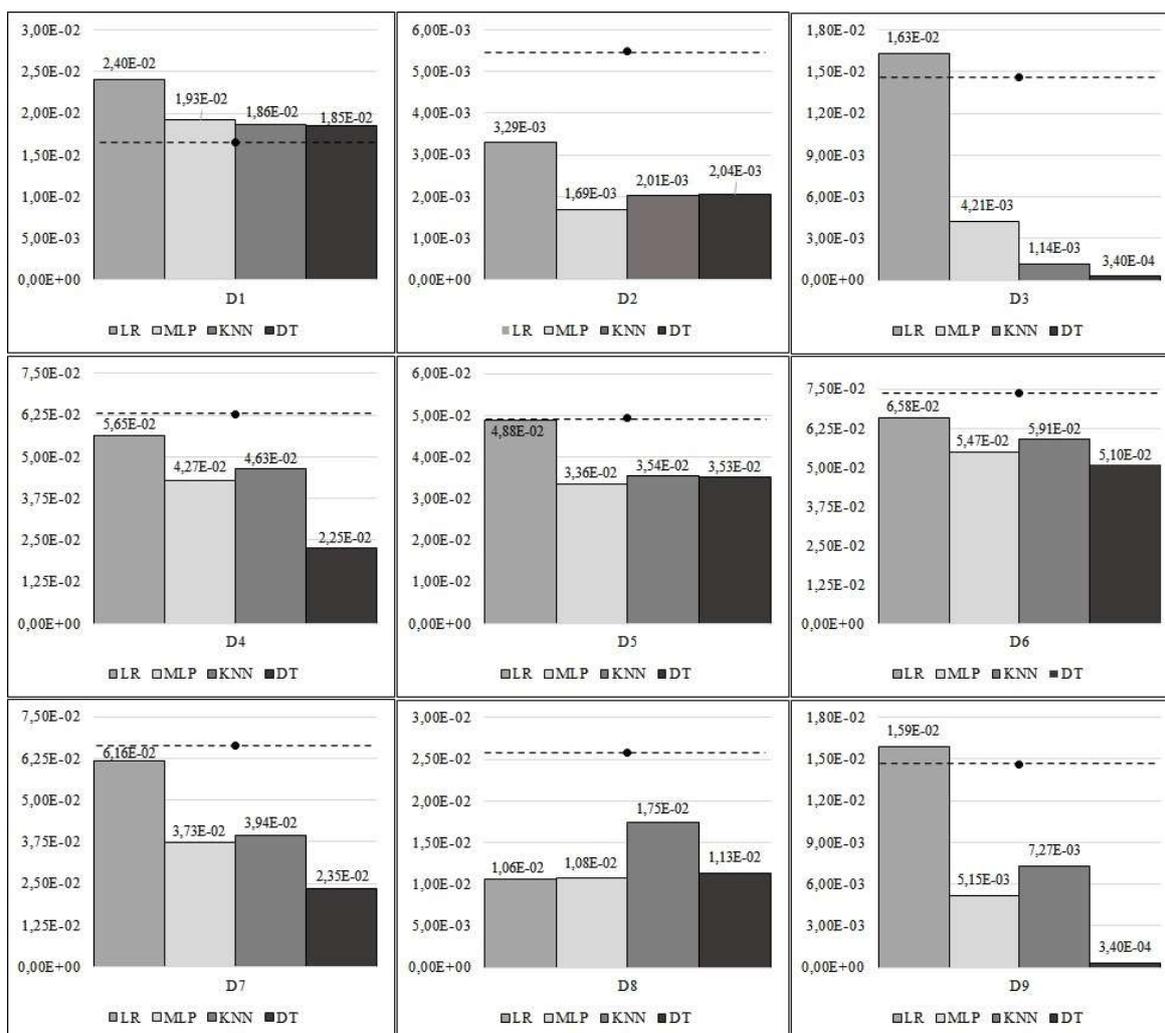
Fonte: Elaborada pelos autores (2023). Nota: Desempenho medido pelo Erro Quadrático Médio entre o valor predito e o valor verdadeiro.

Visualiza-se ainda na Figura 1, que o modelo DT teve o melhor desempenho, não tendo praticamente erros quando aplicado nos diferentes conjuntos de treinamento. Embora a minimização do erro durante o treinamento possa indicar uma boa qualidade de aprendizado, é necessário verificar se não ocorreu *overfitting*, ou seja, se o modelo criado não está superajustado apenas aos exemplos de treinamento, prejudicando assim a capacidade de generalização do modelo.

Diante do exposto, conjuntos de testes compostos por instâncias não vistas durante o treinamento foram submetidos aos modelos preditivos. O desempenho dos modelos nos conjuntos de testes dos nove conjuntos de metadados usados nos experimentos está ilustrado nos gráficos contidos na Figura 2. A linha horizontal pontilhada indica o valor de desempenho do modelo baseline observado em cada conjunto de teste. Apenas no conjunto de metadados D1 todos os modelos tiveram desempenho inferior ao baseline.

Destaca-se que o modelo DT teve o melhor desempenho preditivo na maioria dos conjuntos de teste, sendo que especificamente nos conjuntos D3 e D9, os valores de preço do carbono previstos para as instâncias dos respectivos conjuntos foram muito próximo dos valores verdadeiros. Esses dois conjuntos são constituídos por dados dos municípios amazonenses, quais sejam: quantidades de emissões de GEEs e valores do PIB, respectivamente.

Figura 2 - Desempenho dos Modelos Preditivos nos Conjuntos de Testes



Fonte: Elaborada pelos autores (2023). Nota: Desempenho medido pela Distância Absoluta Média entre o valor predito e o valor verdadeiro.

Considerando que o uso da árvore de decisão apresentou desempenho satisfatório quando alguns conjuntos de metadados foram usados, então decidimos por aplicar simultaneamente várias árvores de decisão para verificar se há melhoria no desempenho preditivo para os demais metadados conforme está detalhado na Seção 4.1.1. Além disso, com a possível relação existente entre os dados amazônicos e o preço do carbono, a Seção 4.1.2 apresenta resultados dos experimentos em que os metadados dos municípios amazonenses foram integrados. E, também, uma vez que o viés da árvore de decisão em selecionar os melhores atributos durante a construção do modelo parece ser relevante no contexto desse problema, novos experimentos computacionais descritos na Seção 4.1.3 foram conduzidos usando técnicas de seleção de atributos.

4.1.1 Modelo *Ensemble*

Quando decisões críticas precisam ser tomadas em uma organização, é aconselhável que sejam feitas por um conselho deliberativo, evitando que a decisão seja monocrática e, assim sendo feito, aumentam-se as chances de sucesso.

De forma similar, em aprendizado de máquina podemos ter o valor de saída como resultado da combinação dos valores fornecidos por diferentes modelos. Dependendo da quantidade e da qualidade dos dados de treinamento pode ser que o resultado produzido por um conjunto de modelos, denominado modelo *ensemble*, tenha um desempenho preditivo melhor do que aquele observado por um modelo simples (Witten & Eibe, 2005).

É possível combinar diferentes árvores de decisão aplicadas para o mesmo problema e ter um modelo preditivo gerado pelo algoritmo floresta aleatória, em que cada árvore modela de forma independente e aleatória uma amostra dos dados originais. Além disso, cada árvore pode encontrar a melhor divisão de um nó a partir de um subconjunto de variáveis preditoras selecionadas aleatoriamente (Cutler *et al.*, 2012).

A Tabela 4 mostra a média e o desvio padrão da diferença absoluta entre o valor verdadeiro e o valor predito tanto pelo modelo de árvore de decisão quanto pelo modelo floresta aleatória em cada conjunto de teste proveniente dos nove conjuntos de metadados investigados nesta pesquisa.

Tabela 4 - Desempenho da Árvore de Decisão e Florestas Aleatórias nos Conjuntos de Teste

ID	ÁRVORE DE DECISÃO	FLORESTAS ALEATÓRIAS
D1	18,46 ± 0,00	18,46 ± 0,00
D2	2,04 ± 0,00	2,04 ± 0,00
D3	0,34 ± 0,00	0,34 ± 0,00
D4	22,33 ± 18,92	22,32 ± 18,67
D5	35,28 ± 3,21	35,27 ± 3,20
D6	51,00 ± 0,00	51,00 ± 0,00
D7	23,52 ± 17,89	23,52 ± 17,89
D8	11,29 ± 0,00	11,29 ± 0,00
D9	0,34 ± 0,00	0,34 ± 0,00

Fonte: Elaborada pelos autores (2023).

Comparando os valores calculados pela Diferença Absoluta Média, os dois modelos tiveram o mesmo desempenho preditivo na maioria dos conjuntos de metadados. Observa-se que o desempenho do modelo induzido pelo algoritmo de AM floresta aleatória foi ligeiramente melhor apenas nos conjuntos de metadados D4 e D5, sendo registrado desempenho igual para os demais conjuntos. Para confirmar se de fato existe alguma diferença significativa nos valores preditos pelos dois modelos para as instâncias de testes nos conjuntos D4 e D5, é recomendável aplicar um teste estatístico. Seguindo os estudos de Demsar (2006), o Teste T foi calculado considerando o valor para a variável $\alpha = 5\%$.

Ao aplicar o Teste T para $\alpha = 5\%$ nos valores fornecidos pelos modelos preditivos nos conjuntos de metadados D4 e D5, obteve-se os valores de probabilidade $p = 0,953$ e $p = 0,964$, respectivamente. Uma vez que estes valores de probabilidade foram maiores que o valor de α , isso indica que, para um nível de

confiança de 95%, não foi possível rejeitar a hipótese nula de que os dois modelos tenham o mesmo desempenho preditivo.

4.1.2 Dados Amazônicos

Com a finalidade de obter um único conjunto de metadados apenas com atributos relativos aos municípios amazonenses, procedeu-se a unificação dos conjuntos de metadados sobre gases de efeito estufa, desmatamento e produto interno bruto. Para isso, foi necessário primeiramente selecionar as instâncias que correspondessem ao mesmo intervalo de tempo em todos os três conjuntos. Após a realização desse procedimento, o conjunto de metadados unificado ficou constituído por 744 instâncias, resultado da quantidade das instâncias de cada um dos 62 municípios do estado do Amazonas relativos ao período de 12 anos (2008 a 2019).

Todos os meta-atributos dos três conjuntos de metadados foram incorporados, totalizando em 29 meta-atributos. Os modelos preditivos foram treinados com esse novo conjunto de metadados, mas antes um subconjunto das instâncias foi separado para a realização de testes dos modelos, seguindo a mesma metodologia já mencionada. O desempenho dos modelos no conjunto de testes foi mensurado pela diferença absoluta média cujo valores estão expostos na Tabela 5. Os resultados apontam que o desempenho preditivo do modelo implementado a partir do algoritmo floresta aleatória foi exatamente igual ao obtido pelo modelo baseado em árvore de decisão. Esse resultado mostra que o uso de várias árvores de decisão não propiciou uma previsão mais precisa do preço do carbono em relação ao uso de apenas uma única árvore.

Tabela 5 - Desempenho dos Modelos nos Conjuntos de Teste com Metadados Amazônicos.

MODELO	MÉDIA	DESVIO-PADRÃO
Baseline	14,53	0,00
Regressão Linear	12,46	1,62
MLP	4,31	1,27
KNN	8,21	3,00
Árvore de Decisão	0,30	0,00
Floresta Aleatória	0,30	0,00

Fonte: Elaborada pelos autores (2023).

4.1.3 Seleção de Atributos

A fim de investigar se os modelos de AM poderiam ser capazes de melhorar ainda mais o seu desempenho preditivo, um processo de seleção de atributos no conjunto de metadados amazônico foi aplicado. Dos 29 meta-atributos oriundos dos 3 conjuntos de metadados, 4 meta-atributos foram excluídos por serem redundantes. Os meta-atributos “Ano” e “Município” estavam presentes nos 3 conjuntos originais, então por não haver necessidade de repetição desses meta-atributos extraiu-se apenas de um dos conjuntos.

A seleção de meta-atributos foi realizada por meio do método *feature_selection.f_regression* disponível na ferramenta *sklearn*. Este método retorna um valor de probabilidade que indica o nível potencial de associação de cada meta-atributo preditivo com a variável alvo (Pedregosa *et al.*, 2011).

Com a execução desse método no conjunto de metadados amazônico, obteve-se os valores de probabilidade que estão expostos na Tabela 6. Diante dos resultados obtidos, estabeleceu-se como critério de seleção de meta-atributos o valor de probabilidade igual a 1,0.

Tabela 6 - Valores de Probabilidade Resultante da Aplicação da Seleção de Atributos nos Dados Amazônicos

CONJUNTO DE META-ATRIBUTOS	META-ATRIBUTO	PROBABILIDADE
GEE AMAZONAS	Ano	2,40x10 ⁻²⁰
	Município	1,00
	Emissao	3,91x10 ⁻¹
PIB AMAZONAS	Mesorregiao	1,00
	Microrregiao	1,00
	Regiao_Geog_Imediata	1,00
	Regiao_Geog_Intermediaria	1,00
	Agropecuaria	2,63x10 ⁻¹
	Industria	8,27x10 ⁻¹
	Servicos	7,99x10 ⁻¹
	Outros_Servicos	5,24x10 ⁻¹
	Bruto_Total	7,85x10 ⁻¹
	Imposto_Liquidos	8,49x10 ⁻¹
	PIB	7,98x10 ⁻¹
DESMATAMENTO AMAZONAS	PIB per capta	7,27x10 ⁻¹
	Latitude	1,00
	Longitude	1,00
	AreaKm2	1,00
	Desmatada	4,04x10 ⁻¹
	Incremento	7,89x10 ⁻³
	Floresta	6,20x10 ⁻¹
	Nuvem	3,60x10 ⁻²
	NaoObservado	3,84x10 ⁻²
	NaoFloresta	1,00
Hidrografia	1,00	

Fonte: Elaborada pelos autores (2023).

A partir do critério adotado para a escolha dos atributos, um conjunto de 10 meta-atributos foi selecionado, originando em um novo conjunto de metadados. O mesmo procedimento de treinamento e teste dos modelos preditivos foi realizado com este novo conjunto de metadados. Os resultados relativos ao desempenho preditivo dos modelos estão dispostos na Tabela 7.

Tabela 7 - Comparação de desempenho dos modelos de AM usando metadados com diferentes conjuntos de atributos

MODELO	TODOS OS ATRIBUTOS		SELEÇÃO DE ATRIBUTOS	
	MÉDIA	DESVIO-PADRÃO	MÉDIA	DESVIO-PADRÃO
Baseline	14,53	0,00	14,53	0,00
Regressão Linear	12,46	1,62	12,53	1,90
MLP	4,31	1,27	4,48	0,44
KNN	8,21	3,00	14,01	0,77
Árvore de Decisão	0,30	0,00	14,53	0,00
Floresta Aleatória	0,30	0,00	14,53	0,19

Fonte: Elaborada pelos autores (2023).

É notório que usando o subconjunto de meta-atributos selecionados não houve melhoria na capacidade preditiva dos modelos. Isso mostra que mesmo selecionando um conjunto constituído pelos melhores atributos ranqueados em relação à sua associação com o atributo alvo, não necessariamente os resultados serão mais promissores. Os melhores atributos considerados de forma isolada não garantem que quando agrupados e usados por modelos de AM irá proporcionar um melhor desempenho preditivo conforme mencionado por Faceli *et al.* (2011).

5 Considerações Finais

As mudanças climáticas têm se agravado em nosso planeta nas últimas décadas e isso têm sido motivos de preocupações constantes por parte de muitos agentes da sociedade, tais como: governos, organizações, empresas e população. Propostas de ações para preservação da natureza são discutidas anualmente em conferências internacionais, onde países ratificam seus compromissos ambientais, dentre os quais está o controle de emissões de gases de efeito estufa - GEEs.

Uma das estratégias atuais para minimizar a emissão de GEEs é por meio do comércio de crédito de carbono. Empresas e países desenvolvidos poderão ter direito a crédito de carbono por meio de investimentos em projetos que incorporam mecanismos de desenvolvimento limpo e que sejam executados em países em desenvolvimento. Uma tonelada de dióxido de carbono que deixa de ser emitida na atmosfera equivale a um crédito de carbono. Desse modo, empresas e países detentoras de créditos de carbono podem negociá-los com outras empresas e países que não conseguem atingir suas metas obrigatórias de redução dos GEEs.

É nesse contexto que o Brasil pode obter alguma vantagem competitiva, já que possui a maior floresta do mundo, a Floresta Amazônica, que tem a capacidade de armazenar bilhões de toneladas de carbono em seu ecossistema. Alguns países sabendo dessa realidade já injetam recursos financeiros nessa região para obter créditos de carbono. No entanto, ainda é possível atrair investimentos de muitos outros países desenvolvidos, tendo em vista a preocupação ecológica atual, refletindo na necessidade da preservação da Floresta Amazônica. Com mais investimentos, certamente deverá haver uma redução do número de focos de queimadas e da área desmatada na região. Além dos benefícios ambientais, os municípios da região também poderão ser favorecidos com um crescimento da economia local acompanhado do desenvolvimento social.

Atualmente, existem muitas pesquisas que buscam prever o preço do carbono a partir de modelos econométricos e modelos baseados em IA. Em nossos experimentos computacionais, modelos de IA baseados em diferentes técnicas de aprendizado de máquina foram implementados para estimar o preço do carbono cotado no principal mercado mundial. A capacidade preditiva dos modelos foi avaliada com o uso de 9 conjuntos de metadados relativos à sustentabilidade em contexto mundial, nacional e regional.

Os resultados experimentais foram mais promissores quando os conjuntos de metadados sobre os municípios do estado do Amazonas foram submetidos ao modelo baseado na técnica baseada em árvore de decisão. Por essa razão, uma combinação de várias árvores de decisão, denominada floresta aleatória (*Random Forest* no original em inglês), foi também avaliada, sendo que no geral o desempenho foi similar ao observado pelo modelo induzido por uma única árvore.

Um experimento adicional foi realizado a partir da união dos conjuntos de dados sobre os aspectos amazônicos (GEE, PIB e desmatamento), no qual foi aplicado um processo de seleção de atributos. Nesse processo, foram escolhidos os atributos que apresentaram uma melhor relação com o valor de saída. Após o processamento de treinamento e teste dos modelos preditivos com esse novo conjunto de metadados, percebeu-se que não houve melhoria na capacidade preditiva dos mesmos. Isso confirma que nem sempre os melhores atributos isoladamente vão proporcionar resultados mais satisfatórios quando usados em um mesmo conjunto.

A possível relação existente entre os dados amazônicos e o preço do carbono no mercado mundial mostra que muitos países têm interesse no que acontece com a Floresta Amazônica. Não há dúvidas de que a preservação da maior floresta do mundo é imprescindível para o bem-estar de todos os seres que vivem na Terra. Por isso, para que a Região Amazônica obtenha vantagens com a geração de créditos de carbono a partir da sua riqueza natural é importante que o Brasil defina políticas públicas sustentáveis e crie os marcos legais para a criação do mercado de carbono no país.

Como propostas de continuidade desta pesquisa, sugerimos o uso de modelos preditivos induzidos por algoritmos de AM para outros conjuntos de dados da região amazônica, como por exemplo: histórico de queimadas e de desmatamentos dos municípios amazonenses. Além disso, dados mais recentes poderiam ser coletados das fontes primárias para atualizar os conjuntos de metadados, de tal modo que possam ser submetidos aos modelos preditivos. E, dessa maneira, prosseguir com a investigação sobre o uso de modelos computacionais para prever o preço do carbono.

Agradecimentos

Os autores agradecem o apoio recebido pelas seguintes instituições: Universidade Federal do Amazonas, Universidade de São Paulo e Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Referências

- Albuquerque, L., Gavioli, L., Margulis, S., Barreto, C., Clemente, F., Goulart, J. & Esposito, S. (2021). *Relatório 2021 - oportunidades para o Brasil em mercado de carbono*. São Paulo: International Chamber of Commerce Brasil e Way Carbon.
- Assis, T. F., & D'Agosto, M. de A. (2020). *Guia para inventário de emissões: gases de efeito estufa nas atividades logísticas* (1st ed.). Rio de Janeiro: Instituto Brasileiro de Transporte Sustentável.
- Brasil. (2017). *Sistema de registro nacional de emissões*. Brasília: Ministério da Ciência, Tecnologia e Inovação. Recuperado de: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/sirene>
- Brasil. (2021). *Resultados do inventário nacional das emissões de gases de efeito estufa por unidade federativa*. Brasília: Ministério da Ciência, Tecnologia e Inovação. Recuperado de: <https://www.gov.br/mcti/pt-br/acompanhe-o-cti/sirene/arquivos/livroresultadoinventario30062021>
- Brasil. (2022). *Decreto nº. 11.075 de 19 de maio de 2022. Estabelece os procedimentos para a elaboração dos Planos Setoriais de Mitigação das Mudanças Climáticas, institui o Sistema Nacional de Redução de Emissões de Gases de Efeito Estufa e altera o Decreto nº 11.003, de 21 de março de 2022*. Diário Oficial da União, 94-A, Brasília. Recuperado de: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/decreto/D11075.htm
- Cebds. (2021). *Mercado de carbono - guia cebds*. Conselho Empresarial Brasileiro para o Desenvolvimento Sustentável. Recuperado de: [https://cebds.org/wp-content/uploads/2021/08/cartilha MercadoDeCarbono 2021.pdf](https://cebds.org/wp-content/uploads/2021/08/cartilhaMercadoDeCarbono2021.pdf)
- Chen, M., Mao, S., & Liu, Y. (2014, abr). Big data: A survey. *Mob. Netw. Appl.*, 19(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Costa, A. B. d., & Soares, M. A. T. (2020, set.). Crise econômica, pandemia e a política de transferência de renda no Brasil. *Cadernos de Ciências Sociais Aplicadas*, 17(30), 65-83. Recuperado de: <https://doi.org/10.22481/ccsa.v17i30.7124>
- Cutler, A., Cutler, D.R., & Stevens, J.R. (2012). Random forests. In: Zhang, C., Ma, Y. (eds) *Ensemble Machine Learning*. Springer, Boston, MA. Recuperado de: https://doi.org/10.1007/978-1-4419-9326-7_5
- Demsar, J. (2006, 01). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7, 1-30. Recuperado de: <https://www.jmlr.org/papers/volume7/demsar06a/demsar06a.pdf>
- Dias, G., Neffa, E., & Tostes, J. G. R. (2020, dez.). Pagamentos por serviços ambientais, mercado de crédito de carbono e as trocas desiguais. *Boletim do Observatório Ambiental Alberto Ribeiro Lamego*, 14(2), 235-254. Recuperado de: <https://doi.org/10.19180/2177-4560.v14n22020p235-254>

- Dormady, N. C. (2014). Carbon auctions, energy markets & market power: An experimental analysis. *Energy Economics*, 44, 468-482. Recuperado de: <https://doi.org/10.1016/j.eneco.2014.03.013>
- Eibel, E., & Pinheiro, R. (2015, out.). Crédito de carbono. *Revista Gestão & Sustentabilidade Ambiental*, 4(2), 588-601. Recuperado de: <https://doi.org/10.19177/rgsa.v4e22015588-601>
- Faceli, K., Lorena, A. C., Gama, J., & de Carvalho, A. C. P. L. F. (2011). *Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro: LTC.
- Fan, J. H., & Todorova, N. (2017). Dynamics of china's carbon prices in the pilot trading phase. *Applied Energy*, 208, 1452-1467. Recuperado de: <https://doi.org/10.1016/j.apenergy.2017.09.007>
- Fearsinde, P. M. (1985). Brazil's amazon forest and the global carbon problem. *Interciência*, 10(4), p.179-186. Recuperado de: https://www.researchgate.net/publication/285026296_Brazil's_Amazon_Forest_and_the_Global_Carbon_Problem
- Galbraith, D., Malhi, Y., Aragão, L. & Baker, T. (2014, abr). The ecosystem dynamics of Amazonian and Andean forests. *Plant Ecology & Diversity*, 7(1-2), 1-6. Recuperado de: <https://doi.org/10.1080/17550874.2013.826744>
- Gerhardt, A., T.; De Souza. (2009). Aspectos teóricos e conceituais. In Gerhardt, T & Silveira D. (Eds.), *Métodos de pesquisa*. p. 13-31. Porto Alegre: UFRGS. Recuperado de: <https://lume.ufrgs.br/handle/10183/213845>
- Gil, A. C. (2010). *Como elaborar projetos de pesquisa*. São Paulo: Atlas.
- ICAP. (2022). *About us*. International Carbon Action Partnership. Recuperado de: <https://icapcarbonaction.com/en/about-us>
- INPE. (2020). *Portal do monitoramento de queimadas e incêndios florestais*. Instituto Nacional de Pesquisas Espaciais. Recuperado de: <https://queimadas.dgi.inpe.br/queimadas/portal>.
- Kaack, L., Donti, P., Strubell, E., Kamiya, G., Creutzig, F. & Rolnick, D. (2022, jun.). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12, 518-527. Recuperado de: <https://doi.org/10.1038/s41558-022-01377-7>
- Kaack, L., Donti, P., Strubell, E. & Rolnick, D. (2020). Artificial intelligence and climate change: Opportunities, considerations, and policy levers to align ai with climate change goals. *Heinrich Böll Stiftung*. Recuperado de: https://eu.boell.org/sites/default/files/2021-04/Artificial%20Intelligence%20and%20Climate%20Change_FINAL14042021.pdf
- Kanda, J. Y., & Carvalho, A. C. P. de L. F. de (2022). Uso da inteligência artificial na previsão do preço do carbono. *Anais do III congresso on-line interacional de sustentabilidade*. Fortaleza, Ceará, Brasil, 3(4), 590-594. Recuperado de: <https://editoraime.com.br/revistas/index.php/rema/issue/view/62/35>
- Kim, S.-K., & Huh, J.-H. (2020, 05). Blockchain of Carbon Trading for UN Sustainable Development Goals. *Sustainability*, 12(10), 1-32. Recuperado de: <https://doi.org/10.3390/su12104021>
- McAfee, A., & Brynjolfsson, E. (2012, out.). Big data: The management revolution. *Harvard business review*, 90(10), 61-67. Recuperado de: https://www.researchgate.net/publication/232279314_Big_Data_The_Management_Revolution

- Mikhailova, I. (2004). Sustentabilidade: evolução dos conceitos teóricos e os problemas da mensuração prática. *Revista Economia e Desenvolvimento*, 16, 22-41. Recuperado de: <https://core.ac.uk/download/pdf/231146376.pdf>
- Minayo, M. C. de S. (2014). *O desafio do conhecimento: pesquisa qualitativa em saúde*. São Paulo: Hucitec.
- Mitchell, T. (1997). *Machine learning*. New York: McGraw Hill.
- Oliveira, A. (2018, 04). Mercado de crédito de carbono voluntário pelas empresas brasileiras: análise sobre a (in)eficiência deste instrumento para o meio ambiente ecologicamente equilibrado. *Direito Ambiental e Sociedade*, 8(1), 83-105. Recuperado de: <https://doi.org/10.18226/22370021.v8.n1.05>
- Passarinho, N. (2021, 13 nov.). Cop 26: Os principais fracassos e vitórias do acordo final da cúpula sobre mudança climática. BBC News Brasil. Recuperado de: <https://www.bbc.com/portuguese/internacional-59274397>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830. Recuperado de: <https://jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>
- Pinheiro, A., Barreto, C., Assumpção, F., Albuquerque, L., Gavioli, L., Garcia, L. & Marques, R. (2022). *Relatório 2022 - oportunidades para o Brasil em mercado de carbono*. São Paulo: International Chamber of Commerce Brasil e Way Carbon.
- Piva, L. R. d. O., Sanquetta, C. R., Wojciechowski, J., & Corte, A. P. D. (2021). Estoques de biomassa e carbono na Amazônia brasileira: uma nova abordagem. *Revista em Agronegócio e Meio Ambiente*, 14, 1-20. Recuperado de: <https://periodicos.unicesumar.edu.br/index.php/rama/article/view/8820>
- Rossum, G. V., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- SEEG. (2019). *Entenda as estimativas*. Eco Stage. Recuperado de: <https://seeg.eco.br/entenda-as-estimativas>
- SEEG. (2022). *Sistema de estimativa de emissões de gases de efeito estufa*. Eco Stage. Recuperado de: https://plataforma.seeg.eco.br/total_emission
- Silveira, D., & Córdova, F. A. (2009). A pesquisa científica. In Gerhardt, T & Silveira D. (Eds.), *Métodos de pesquisa* (p. 33-44). Porto Alegre: UFRGS. Recuperado de: <https://lume.ufrgs.br/handle/10183/213838>
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining*. Boston: Addison Wesley.
- UFAM. (2019). Atlas dos objetivos de desenvolvimento sustentável no Amazonas. Universidade Federal do Amazonas. Recuperado de: <https://www.atlasodsamazonas.ufam.edu.br/>
- Unfccc. (2018). *Achievements of the clean development mechanism: Harnessing incentive for climate action*. United Nations Framework Convention on Climate Change. Recuperado de: https://unfccc.int/sites/default/files/resource/UNFCCC_CDM_report_2018.pdf

- Whittington, J. (2016, 05). Carbon as a metric of the human. *PoLAR: Political and Legal Anthropology Review*, 39(1), 46-63. Recuperado de: <https://doi.org/doi:10.1111/plar.12130>
- Witten, I. H., & Eibe, F. (2005). *Data mining: practical machine learning tools and techniques*. San Francisco, CA: Morgan Kaufmann.
- WorldBank. (2022). *State and trends of carbon pricing 2022*. Washington, DC: The World Bank. Recuperado de: <https://openknowledge.worldbank.org/handle/10986/37455>
- Wynsberghe, A. V. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI Ethics*, 1, 213-218. Recuperado de: <https://link.springer.com/article/10.1007/s43681-021-00043-6>
- Zhang, J., Li, D., Hao, Y., & Tan, Z. (2018). A hybrid model using signal processing technology, econometric models and neural network for carbon spot price forecasting. *Journal of Cleaner Production*, 204, 958-964. <https://doi.org/10.1016/j.jclepro.2018.09.071>
- Zhang, Z., & Huisingh, D. (2017). Carbon dioxide storage schemes: Technology, assessment and deployment. *Journal of Cleaner Production*, 142, Part 2, p.1055-1064. Recuperado de: <https://doi.org/10.1016/j.jclepro.2016.06.199>
- Zhou, J., Huo, X., Xu, X., & Li, Y. (2019, 03). Forecasting the carbon price using extreme-point symmetric mode decomposition and extreme learning machine optimized by the grey wolf optimizer algorithm. *Energies*, 12(5), 1-22. Recuperado de: <https://doi.org/10.3390/en12050950>

Dados dos autores:

Jorge Yoshio Kanda

ORCID: <http://orcid.org/0000-0001-6710-2157>

Doutorado em Ciências da Computação e Matemática Computacional pela Universidade de São Paulo, USP, Brasil. Professor Associado pela Universidade Federal do Amazonas, UFAM, Brasil. E-mail: jkanda@ufam.edu.br

André Carlos Ponce de Leon Ferreira de Carvalho

ORCID: <http://orcid.org/0000-0002-4765-6459>

Pós-doutorado pela University of Kent of Canterbury, UKC, Grã-Bretanha. Pós-doutorado pela Universidade do Porto, Porto, Portugal. Livre-docência pela Universidade de São Paulo, USP, Brasil. Professor pela Universidade de São Paulo, USP, Brasil. E-mail: andre@icmc.usp.br

Como citar este artigo:

Kanda, J. Y. & Carvalho, A. C. P. de L. F. de (2023). Previsão do preço do carbono por modelos de aprendizado de máquina. *AOS - Amazônia, Organizações e Sustentabilidade*, 12(2). <http://dx.doi.org/10.17648/aos.v12i2.2916>