



USO DE INTELIGÊNCIA ARTIFICIAL NA PREVISÃO DO PREÇO DO CARBONO

JORGE YOSHIO KANDA; ANDRÉ CARLOS PONCE DE LEON FERREIRA DE
CARVALHO

RESUMO

O crédito de carbono equivale a uma tonelada de dióxido de carbono não emitida na atmosfera. O comércio de crédito de carbono é um mecanismo atual utilizado por muitos países para tentar impedir o aumento das emissões dos gases de efeito estufa (GEE) e conter as mudanças climáticas na Terra. Esta pesquisa propõe o uso de modelos de regressão baseados em inteligência artificial, mais especificamente, em aprendizado de máquina para prever o preço do carbono. Para isso, foram coletados dados de emissão de GEE relativo ao estado do Amazonas e dados do preço do carbono praticado nos principais mercados mundiais. Os modelos preditivos foram treinados usando a metodologia de validação cruzada em dez pastas. Os resultados preliminares dos experimentos computacionais mostram que os valores relativos à emissão de GEE do estado do Amazonas tem uma forte relação com os preços praticados no mercado de carbono da União Europeia. Isso confirma o fato de que o mundo está preocupado com as ações que afetam diretamente a Floresta Amazônica. A riqueza natural desta floresta pode ser uma grande oportunidade para preservar a natureza e impulsionar o crescimento econômico da Região Amazônica com a geração de créditos de carbono por meio de criação de projetos mecanismos de desenvolvimento limpo.

Palavras-chave: Desenvolvimento Sustentável; Mercado de Carbono; Gases de Efeito Estufa; Amazonas; Aprendizado de Máquina.

ABSTRACT

The carbon credit is equivalent to one ton of carbon dioxide not emitted into the atmosphere. Carbon credit trading is a current mechanism used by many countries to try to stop the increase in greenhouse gases (GHG) emissions and contain climate change on Earth. This research proposes the use of regression models based on artificial intelligence, more specifically, on machine learning to predict the carbon price. GHG emission datasets recorded in the state of Amazonas and a carbon price dataset in the main world markets were collected. The predictive models were trained using the 10-folder cross-validation methodology. The preliminary results of the computational experiments show that the values related to the GHG emission of the state of Amazonas have a strong relationship with the prices practiced in the carbon market of the European Union. This confirms the fact that the world is concerned about actions that directly affect the Amazon rainforest. The natural wealth of this forest can be a great opportunity to preserve nature and boost the economic growth of the Amazon Region with the generation of carbon credits through the creation of clean development mechanisms projects.

Key Words: Sustainable Development; Carbon Market; Greenhouse Gases; Amazon; Machine Learning.

1 INTRODUÇÃO

A recuperação da economia brasileira após as consequências provocadas pela pandemia da COVID-19 deveria estar alinhada à promoção de um desenvolvimento sustentável. Neste contexto, é importante que sejam estabelecidos mecanismos de regulação e instrumentos de precificação dos créditos de carbono para reduzir as emissões de gases de efeito estufa (GEE). Cada crédito de carbono corresponde a uma tonelada de dióxido de carbono não emitida na atmosfera. Empresas e países em desenvolvimento que investem em projetos de Mecanismo de Desenvolvimento Limpo (MDL) têm direito a créditos de carbono, e podem negociá-los com empresas e países desenvolvidos que não conseguem alcançar as suas metas obrigatórias de redução da emissão desses gases (EIBEL e PINHEIRO, 2015).

Albuquerque et al. (2021) mencionam que as principais oportunidades de geração de créditos de carbono no Brasil ocorrem nos setores: agropecuário, florestas, energia, transportes e indústria. Nesses setores, as mitigações de emissões de GEE podem proporcionar uma nova economia para o país, sendo que o processo de reflorestamento por meio do manejo e restauração florestal é a melhor forma para mitigar os efeitos das emissões (CEBDS, 2021).

A Floresta Amazônica tem uma grande importância para o Planeta Terra, pois armazena bilhões de toneladas de carbono em sua biomassa, além de abrigar uma enorme diversidade de árvores e plantas, com inúmeras espécies que fornecem uma série de serviços ecossistêmicos de inestimável valor local, regional e global (GALBRAITH et al., 2014).

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 poderiam contribuir para o aquecimento da Terra entre 1,5 e 4,5°C antes do ano de 2050. Todavia, ele reconheceu a falta de precisão em sua projeção sobre esse problema devido à escassez de dados confiáveis a respeito da taxa de devastação da Amazônia Brasileira e de carbono liberado nessa região.

Atualmente, dados históricos sobre as mudanças climáticas podem ser encontrados em diferentes fontes, como por exemplo: Sistema de Registro Nacional e Emissões (MCTI, 2017), Sistema de Estimativa de Emissões de Gases de Efeito Estufa por Município (SEEG, 2022), além dos indicadores dos municípios amazonenses em relação a cada um dos 17 Objetivos de Desenvolvimento Sustentável definidos pela Organização das Nações Unidas (UFAM, 2019).

Dados produzidos continuamente em grande volume, dispostos em diferentes formatos e originados de fontes diversas constituem um Big Data (CHEN, MAO e LIU, 2014). É bastante desafiador proceder uma análise das informações contidas em um Big Data de modo que possam contribuir para um gestor tomar as melhores decisões (MCAFEE e BRYNJOLFSSON, 2012). Por isso, o uso adequado de recursos computacionais para armazenar, processar, analisar e encontrar possíveis relações entre os dados disponíveis podem ser de grande valia na tomada de decisões para os gestores que desejam definir políticas públicas na questão ambiental.

Extraír informações úteis a partir de um grande volume de dados implica identificar possíveis padrões existentes no conjunto de dados. Essa ação pode ser conduzida por meio de inteligência artificial usando algoritmos implementados a partir de técnicas de aprendizado de máquina (AM) para extrair informações em conjunto de dados, que contém uma grande quantidade de instâncias (TAN, STEIBACH, KUMAR, 2006).

Diante desse contexto, o principal objetivo desta pesquisa consiste em criar modelos computacionais que sejam capazes de usar dados de emissão de GEE registrados no estado do Amazonas para prever o preço do crédito de carbono. Para isso, dados de fontes oficiais devem ser coletados e formatados adequadamente para serem usados pelos algoritmos de AM, resultando em modelos preditivos de regressão. Além disso, deve ser mensurada a capacidade preditiva dos modelos por meio de medidas apropriadas.

2 MATERIAIS E MÉTODOS

Um conjunto de dados sobre a emissão de GEE relativos aos municípios do estado do Amazonas registrados nos anos de 2000 a 2019 foi extraído da plataforma SEEG Municípios, tendo sido selecionados todos os setores responsáveis pela emissão (SEEG, 2022). Um arquivo contendo o preço de carbono registrado no mercado da União Europeia foi baixado da página Web da *International Carbon Action Partnership* (ICAP, 2022).

Os dados coletados passaram primeiramente por uma análise de valores incorretos, inconsistentes, ausentes e duplicados. Além disso, os valores das instâncias foram normalizados no intervalo [0,1] por atributo. Dessa forma, seguindo as recomendações de Faceli *et al.* (2011), cada conjunto de dados foi transformado em um conjunto de metadados com a finalidade de melhorar a capacidade preditiva dos modelos de AM em prever corretamente o valor para dados não vistos durante o treinamento dos modelos.

Para treinar os modelos preditivos foi adotada a metodologia de validação cruzada em 10 pastas. Todas as instâncias do conjunto de dados são igualmente distribuídas em 10 pastas e 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, STEIBACH, KUMAR, 2006). Para mensurar o desempenho dos modelos nos conjuntos de treinamento e nos conjuntos de testes pode ser calculado o Erro Quadrático Médio e a Distância Absoluta Média, respectivamente (FACELI *et al.* 2011).

Os modelos computacionais foram construídos a partir de técnicas de regressão linear (RL), rede neural artificial (RNA), *k-Nearest Neighbors* (k-NN) e árvore de decisão (AD), que estão dentre as mais utilizadas pela comunidade científica da área de Aprendizado de Máquina (ZHOU; YU; YUAN, 2018). Durante a implementação dos modelos foi usado o software *Python* versão 3.10.6 (PYTHON, 2022) e a ferramenta *scikit-learn* versão 1.1.2, da qual foram utilizados os pacotes *LinearRegression*, *MLPRegressor*, *KNeighborsRegressor* e *DecisionTreeRegressor*, respectivamente (PEDREGOSA *et al.* 2011). Todos os algoritmos foram executados em um laptop com processador Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz 2.71 GHz, 4 GB de memória RAM e Sistema Operacional Windows de 64 bits. Os conjuntos de metadados estão disponíveis em: <https://github.com/Kanda2023?tab=repositories>.

3 RESULTADOS E DISCUSSÃO

O desempenho registrado na fase de treinamento dos modelos preditivos pode ser conferido na Figura 1. Percebe-se que o modelo baseado em árvore de decisão obteve o melhor desempenho, em que o erro quadrático médio foi igual a zero. Esse resultado, provavelmente, ocorreu em decorrência de que durante o treinamento o algoritmo procura escolher o melhor atributo que mapeia os valores de entrada com os valores de saída, sendo desse modo criadas as ramificações da árvore de decisão.

Com a finalidade de simular um conjunto de instâncias novas para avaliar a capacidade preditiva dos modelos induzidos, as instâncias mais recentes dos últimos 12 meses de cada conjunto de metadados foram separadas e não utilizadas no treinamento dos modelos. O conjunto dessas instâncias, denominado conjunto de teste, foi submetido a cada um dos modelos de regressão induzidos por diferentes técnicas de AM. Cada modelo processou para cada valor de entrada (instância descrita pelo conjunto de valores dos meta-atributos) um valor de saída (preço do carbono no mercado europeu). A Tabela 1 mostra os valores da diferença absoluta média entre o preço do carbono predito pelos modelos de AM e o preço verdadeiro para as instâncias do conjunto de teste.

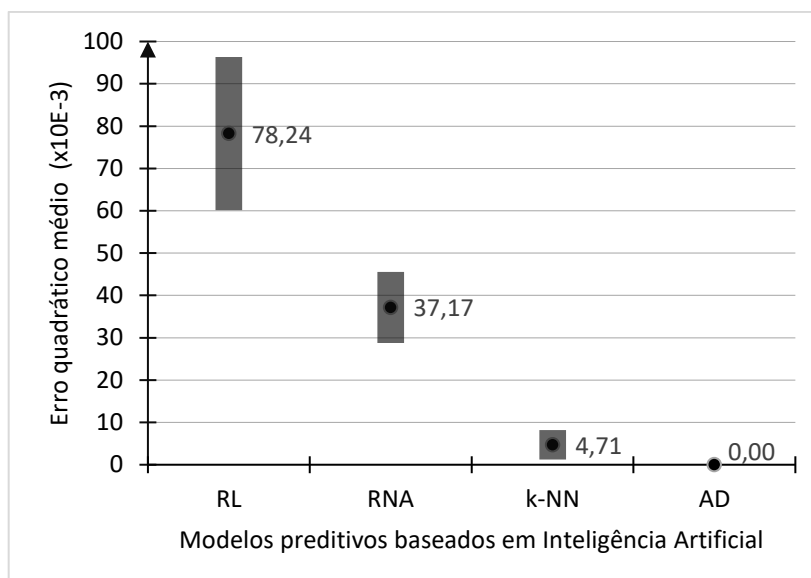


Figura 1. Desempenho dos modelos preditivos no conjunto de treinamento,
Fonte: Autoria Própria (2023).

Tabela 1. Diferença absoluta média entre o preço do carbono predito e o preço verdadeiro.

Modelo Preditivo	Diferença Absoluta Média
Regressão Linear	13,94 ± 0,21
Rede Neural Artificial	4,36 ± 0,60
<i>k-Nearest Neighbors</i>	1,06 ± 2,40
Árvore de Decisão	0,30 ± 0,00

Fonte: Autoria Própria (2023).

O modelo preditivo baseado em árvore de decisão apresentou novamente uma melhor capacidade preditiva em relação aos outros modelos investigados durante os experimentos computacionais. Esse resultado pode ser um indício de que o uso de modelos computacionais baseados em inteligência artificial, mais especificamente, em aprendizado de máquina pode ser bastante promissor para prever o preço do carbono praticado nos mercados mundiais.

4 CONCLUSÃO

As mudanças climáticas que ocorrem atualmente no Planeta Terra são decorrentes de vários fatores dentre os quais está o aumento da emissão dos gases de efeito estufa. Um mecanismo adotado por vários países para controlar essa emissão está no comércio de crédito de carbono. Não há dúvidas de que a Floresta Amazônia tem uma importância mundial no controle das mudanças climáticas e, por isso resultados de pesquisas relacionadas à essa região podem contribuir na criação de conhecimentos que nos permita ter um mundo mais saudável.

Os resultados preliminares desta pesquisa, que usa modelos de inteligência artificial baseado em aprendizado de máquina para prever o preço de carbono, indicam a existência de uma correlação entre os valores de emissão de gases de efeito estufa registrados no estado do Amazonas e o valor do preço do crédito de carbono na União Europeia. Sendo assim, o estado do Amazonas pode ser capaz de atrair investimentos para a geração de créditos de carbono, ajudando na preservação da natureza e contribuindo para o crescimento econômico da região.

Para a continuidade desta pesquisa, sugere-se usar conjuntos de dados de emissão de gases de efeito estufa em âmbito nacional e mundial para verificar se os modelos preditivos continuarão com uma elevada capacidade preditiva. Além disso, podem ser usados outros conjuntos de dados ambientais como desmatamento e queimadas e submetê-los aos modelos.

AGRADECIMENTOS

À Universidade de São Paulo - USP, com a qual os autores possuem vínculos, sendo que o primeiro está no Programa de Pós-Doutorado do Instituto de Ciências Matemáticas e de Computação - ICMC e o segundo é professor livre-docente do ICMC.

À Universidade Federal do Amazonas – UFAM, instituição na qual o primeiro autor é Professor Associado Nível II.

REFERÊNCIAS

ALBUQUERQUE, L. et al. **Relatório 2021 - Oportunidades para o Brasil em Mercado de Carbono**. São Paulo: International Chamber of Commerce Brasil e WayCarbon, 2021.

CEBDS - Conselho Empresarial Brasileiro para o Desenvolvimento Sustentável. **Mercado de Carbono – GUIA CEBDS. 2021**. Disponível em: https://cebds.org/wpcontent/uploads/2021/08/CARTILHA_MercadoDeCarbono_2021.pdf. Acesso em: 15 dez. 2021.

CHEN, M.; MAO, S.; LIU, Y. Big data: A survey. **Mob. Netw. Appl.**, Springer-Verlag, Berlin, Heidelberg, v. 19, n. 2, p. 171–209, 2014. ISSN 1383-469X.

EIBEL, E.; PINHEIRO, R. Crédito de carbono. **Revista Gestão & Sustentabilidade Ambiental**, Florianópolis, v. 4, n. 2, p. 588–601, 2015.

FACELI, K.; LORENA, A.; GAMA, J.; DE CARVALHO, A. **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011.

FEARSINDE, P. M. Brazil's amazon forest and the global carbon problem. **Interciência**, v. 10, n. 4, p. 179–186, 1985.

GALBRAITH, D. MALHI, Y.; ARAGÃO, L.; BAKER, T. The ecosystem dynamics of amazonian and andean forests. **Plant Ecology & Diversity**, v. 7, n. 1-2, p. 1–6, abr. 2014.

ICAP - International Carbon Action Partnership. **About us**. 2022. Disponível em: <https://icapcarbonaction.com/en/about-us>. Acesso em: 27 jul. 2022.

MCAFEE, A.; BRYNJOLFSSON, E. Big data: The management revolution. **Harvard business review**, v. 90, p. 61–67, out. 2012.

MCTI - Ministério da Ciência, Tecnologia e Inovação. **Sistema de Registro Nacional de Emissões**. Brasília, 2017. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/sirene>. Acesso em: 01 jun. 2022.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PYTHON. **Python 3.10.6**. 2022. Disponível em: <https://www.python.org/downloads/release/python-3106/>. Acesso em: 22 ago. 2022.

SEEG. **Sistema de Estimativa de Emissões de Gases de Efeito Estufa**. 2022. Disponível em: <https://plataforma.seeg.eco.br/cities/statistics>. Acesso em: 16 dez. 2022.

TAN, P.-N.; STEINBACH, M.; KUMAR, V. **Introduction to Data Mining**. Boston:

Addison Wesley, 2006.

UFAM. **Atlas dos Objetivos de Desenvolvimento Sustentável no Amazonas**. 2019.
Disponível em: <https://www.atlasodsamazonas.ufam.edu.br/>. Acesso em: 17 maio 2022.

ZHOU, J.; YU, X.; YUAN, X. Predicting the carbon price sequence in the shenzhen emissions exchange using a multiscale ensemble forecasting model based on ensemble empirical mode decomposition. **Energies**, v. 11, n. 1907, p. 1–17, 2018.