

UNIVERSIDADE PRESBITERIANA MACKENZIE

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

Previsão de Emissões de CO₂ e Otimização do Uso de Energias Renováveis: Uma
Abordagem com Machine Learning

São Paulo

2024

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

Previsão de Emissões de CO₂ e Otimização do Uso de Energias Renováveis: Uma
Abordagem com Machine Learning

Trabalho Aplicando conhecimento para entrega no
Moodle referente ao conteúdo 1 de aprendizagem
do componente curricular Ciência, tecnologia e
sociedade;

ORIENTADOR: Prof. GUSTAVO SCALABRINI SAMPAIO

São Paulo

2024

THAINA VIEIRA DOS SANTOS

VINICIUS CAUMO SEGATTO

Previsão de Emissões de CO₂ e Otimização do Uso de Energias Renováveis: Uma
Abordagem com Machine Learning

Trabalho Aplicando conhecimento para entrega no
Moodle referente ao conteúdo 1 de aprendizagem
do componente curricular Ciência, tecnologia e
sociedade;

Aprovado em

BANCA EXAMINADORA

Prof. GUSTAVO SCALABRINI SAMPAIO

Universidade Presbiteriana Mackenzie

SUMÁRIO

1. INTRODUÇÃO

1.1 CONTEXTO

1.2 MOTIVAÇÃO E JUSTIFICATIVA

1.3 REFERENCIAL TEÓRICO

1.4 PIPELINE DO PROJETO

2. OBJETIVO

2.1 CRONOGRAMA

3. DESENVOLVIMENTO

4. REFERÊNCIAS BIBLIOGRÁFICAS

5. APÊNDICES

1. INTRODUÇÃO

1.1 CONTEXTO

A transição para uma matriz energética mais limpa e sustentável é um dos maiores desafios do século XXI. De acordo com uma matéria da *BBC News*, a crise climática tem se manifestado de diversas formas, a seca na Amazônia é um exemplo alarmante de seus impactos⁽¹⁾. A região amazônica, conhecida como o "pulmão do mundo", desempenha um papel crucial na regulação do clima global. A perda de sua biodiversidade e a redução de sua capacidade de absorver carbono contribuem significativamente para o aumento das temperaturas e a intensificação de eventos climáticos extremos. A recente revisão da meta climática brasileira, conforme anunciado pelo governo federal, demonstra um compromisso renovado com a luta contra as mudanças climáticas. Ao retomar a meta estabelecida no Acordo de Paris⁽¹⁾, o país sinaliza sua intenção de reduzir significativamente as emissões de gases de efeito estufa até 2025 e 2030.

A relação entre as emissões de CO₂ e a seca na Amazônia é complexa e multifacetada. O aumento das concentrações de gases de efeito estufa na atmosfera, provenientes principalmente da queima de combustíveis fósseis, intensifica o efeito estufa, elevando as temperaturas globais e alterando os padrões de precipitação. A seca prolongada na Amazônia, por sua vez, reduz a capacidade da floresta de absorver CO₂, gerando um ciclo vicioso que acelera o aquecimento global.

A transição para uma matriz energética mais limpa, baseada em fontes renováveis, é fundamental para mitigar os efeitos das mudanças climáticas e reduzir a frequência e a intensidade de eventos extremos como a seca na Amazônia. Ao substituir os combustíveis fósseis por fontes renováveis, como a energia solar, eólica e hidrelétrica, reduzimos as emissões de gases de efeito estufa e diminuímos a pressão sobre os ecossistemas.

Nosso estudo busca contribuir para o entendimento da relação entre as emissões de CO₂, o uso de energias renováveis e os eventos climáticos extremos. Ao quantificar o

impacto das energias renováveis na redução das emissões e ao identificar os fatores que influenciam a vulnerabilidade dos ecossistemas às mudanças climáticas, podemos desenvolver políticas públicas e estratégias de mitigação mais eficazes.

1.2 MOTIVAÇÃO E JUSTIFICATIVA

A previsão de emissões de CO₂ e a otimização do uso de energias renováveis, com foco no Brasil, representa uma escolha de tema extremamente relevante e estratégica por diversos motivos. O Brasil possui uma matriz energética diversificada, com um grande potencial para o desenvolvimento de fontes renováveis como a hidrelétrica, eólica e solar. O projeto tem o potencial de gerar novas soluções concretas para os desafios relacionados à energia e ao clima. Ao desenvolver modelos precisos de previsão e otimização, poderemos auxiliar na tomada de decisões mais informadas por parte de governos, empresas e sociedade em geral. O tema também abrange diversas áreas do conhecimento, como ciência da computação, engenharia, economia e ciências ambientais. Essa característica multidisciplinar permite a formação de equipes de trabalho mais diversificadas e a geração de soluções mais completas e inovadoras.

1.3 REFERENCIAL TEÓRICO

A produção e o consumo de energia, desde a era industrial, estão intrinsecamente ligados às emissões de gases de efeito estufa (GEE). O dióxido de carbono (CO₂), principal GEE, é um subproduto da queima de combustíveis fósseis, utilizados em larga escala para gerar energia. A permanência do CO₂ na atmosfera por longos períodos contribui para o aumento do efeito estufa e as consequentes mudanças climáticas. A compreensão dessa relação é fundamental para a elaboração de políticas e estratégias que visam mitigar os impactos ambientais da atividade humana⁽⁴⁾.

O Brasil, com sua matriz energética diversificada e rica em fontes renováveis, possui um grande potencial para gerar créditos de carbono e participar ativamente do mercado internacional. Segundo o The World Bank (2021)⁽⁴⁾, o país poderia se beneficiar significativamente da comercialização desses créditos. O setor de transportes, em particular, com o uso crescente do etanol hidratado, apresenta uma oportunidade promissora para a geração de créditos de carbono. No entanto, a falta de uma

regulamentação clara e a ausência de um mercado de carbono consolidado no país ainda são desafios a serem superados.

Paralelamente, a previsão da demanda por combustíveis, como a gasolina, é fundamental para a tomada de decisões estratégicas no setor energético. Estudos como o de Santos Jr. (2018)⁽⁴⁾ demonstram a complexidade dessa tarefa, com diferentes modelos estatísticos sendo utilizados para tentar capturar as nuances do comportamento dos consumidores. A escolha do modelo mais adequado depende de diversos fatores, como a disponibilidade de dados, o período de análise e os objetivos da pesquisa.

De acordo com o artigo "Aplicação de Machine Learning para Previsão das Emissões de CO₂ Produzidas pela Produção de"⁽⁵⁾, de Francisco Soares, o uso de técnicas de Machine Learning (ML) para prever as emissões de CO₂ é uma área em constante evolução. O artigo apresenta uma revisão da literatura sobre o tema, bem como uma aplicação prática de ML para a previsão das emissões de CO₂ de uma empresa de cimento. O artigo apresenta uma aplicação prática de ML para a previsão das emissões de CO₂ de uma empresa de cimento. O autor utiliza um conjunto de dados de 10 anos para treinar um modelo de regressão linear. O modelo é então utilizado para prever as emissões de CO₂ para os próximos 5 anos. Os resultados mostram que o modelo é capaz de prever as emissões de CO₂ com uma boa precisão. O artigo conclui que o uso de ML para a previsão das emissões de CO₂ é uma área promissora. No entanto, ainda há muitos desafios a serem enfrentados, como a falta de dados históricos e a complexidade dos processos de produção.

1.4 PIPELINE DO PROJETO

Uma pipeline representa um fluxo de trabalho sequencial que guia o desenvolvimento do projeto desde a concepção até a conclusão. Ela serve como um roadmap, delineando as etapas e as dependências entre elas. É importante ressaltar que uma pipeline é dinâmica e pode ser adaptada ao longo do desenvolvimento do projeto, conforme novas informações surgem ou as prioridades mudam.

Proposta de Pipeline para o Projeto

1) Definição do Problema e Objetivos:

- Clareza: Definir com precisão o problema que se busca resolver: qual é o objetivo principal da previsão das emissões de CO₂?
- Alcance: Estabelecer o escopo do projeto, delimitando os limites da análise.
- Metas: Definir métricas de sucesso para avaliar o desempenho do modelo.

2) Coleta e Preparação dos Dados:

- Identificação das fontes: Identificar as fontes de dados relevantes para a análise (dados históricos de produção, consumo de energia, emissões, etc.).
- Coleta: Coletar os dados e garantir a sua qualidade.
- Limpeza: Limpar os dados, removendo outliers, inconsistências e valores faltantes.
- Transformação: Transformar os dados para um formato adequado para a análise (normalização, criação de novas variáveis, etc.).
- Exploração: Analisar os dados exploratoriamente para identificar padrões e relações.

3) Seleção e Engenharia de Features:

- Seleção: Escolher as variáveis mais relevantes para a previsão das emissões de CO₂.

- Engenharia: Criar novas variáveis a partir das existentes, se necessário, para melhorar o desempenho do modelo.

4) Escolha do Modelo de Machine Learning:

- Revisão da literatura: Pesquisar os modelos de ML mais adequados para a previsão de séries temporais e para o problema específico.
- Experimentação: Testar diferentes modelos (regressão linear, árvores de decisão, redes neurais, etc.) e comparar seus desempenhos.

5) Treinamento e Validação do Modelo:

- Divisão dos dados: Dividir os dados em conjuntos de treinamento, validação e teste.
- Treinamento: Treinar o modelo escolhido utilizando o conjunto de treinamento.
- Validação: Avaliar o desempenho do modelo no conjunto de validação e ajustar os hiperparâmetros, se necessário.

6) Avaliação do Modelo:

- Métricas de desempenho: Avaliar o desempenho do modelo utilizando métricas adequadas (MAE, RMSE, MAPE, etc.).
- Interpretação: Interpretar os resultados do modelo e analisar a importância das diferentes variáveis.

2. OBJETIVO

Desenvolver um modelo preditivo robusto e preciso para estimar as emissões de CO₂ no Brasil, com o objetivo de otimizar a utilização de fontes de energia renovável e auxiliar na tomada de decisões estratégicas para a transição energética.

Objetivos Específicos Relacionados ao Desenvolvimento do Projeto:

1. Coleta e Preparação de Dados:

- Identificar e coletar dados relevantes de diversas fontes, incluindo bases de dados governamentais, agências internacionais e estudos acadêmicos.
- Realizar a limpeza e o tratamento dos dados, garantindo a qualidade e a consistência das informações.
- Criar um banco de dados integrado e estruturado para facilitar a análise e o desenvolvimento dos modelos.

2. Engenharia de Features:

- Selecionar e criar variáveis relevantes para a previsão das emissões de CO₂, considerando fatores como produção industrial, consumo energético, atividades agrícolas, desmatamento e variáveis climáticas.
- Realizar a engenharia de features para extrair informações relevantes dos dados e melhorar a performance dos modelos.

3. Modelagem Preditiva:

- Experimentar diferentes algoritmos de machine learning para a construção dos modelos preditivos.
- Avaliar a performance dos modelos utilizando métricas adequadas, como RMSE (Root Mean Squared Error), MAE (Mean Absolute Error) e R².
- Selecionar o modelo com melhor desempenho para a previsão das emissões de CO₂.

4. Análise de Cenários:

- Validar os modelos utilizando dados históricos e comparando os resultados com previsões anteriores.
- Refinar os modelos de forma contínua, incorporando novos dados e informações para melhorar a precisão das previsões.
- Desenvolvimento de Ferramentas e Visualizações:

Ao longo do desenvolvimento do projeto, é fundamental realizarmos uma avaliação contínua dos resultados e ajustar a metodologia conforme necessário.

2.1 CRONOGRAMA

- Flexibilização: A inclusão de reuniões semanais ou quinzenais permite um acompanhamento mais próximo do progresso do projeto e a identificação de possíveis problemas em tempo hábil.
- Detalhamento da Etapa 3: A sugestão de detalhar a etapa 3 com a especificação de modelos e métricas é crucial para garantir que a equipe esteja trabalhando em direção aos mesmos objetivos.
- Iteratividade: O destaque para a natureza iterativa do processo de desenvolvimento de modelos de machine learning é fundamental para que a equipe esteja preparada para realizar ajustes e refinamentos ao longo do projeto.

Etapa	Data de Início	Data de Término	Checkpoint	Entregáveis	Observações
Definição do Projeto e equipe	06/09	30/09	22/08	Conforme descrito	Reunião semanal para acompanhar o progresso.
Referencial teórico e cronograma	01/10	15/10	19/09	Conforme descrito	Iniciar a coleta e preparação dos dados.
Implementação parcial	16/10	28/10	17/10	Conforme descrito	Explorar diferentes modelos de machine learning.
Implementação e entrega final	29/10	18/11	-	Relatório final com resultados,	Incluir análise de sensibilidade

				código e apresentaçã o.	e incertezas.
--	--	--	--	-------------------------------	---------------

3. DESENVOLVIMENTO

3.1 DESCRIÇÃO DA BASE DE DADOS

1. Dados de emissões são provenientes do Climate Watch Historical GHG Emissions (Emissões Históricas de Gases de Efeito Estufa) de 1990 a 2020⁽²⁾

- **Fonte dos dados:** Os dados sobre as emissões de gases de efeito estufa (GEE) utilizados neste conjunto de dados foram obtidos do Climate Watch, uma plataforma online que oferece dados, visualizações e recursos sobre o clima.
- **Período:** Os dados abrangem um período de 30 anos, desde 1990 até 2020.
- **Tipo de dado:** O conjunto de dados se concentra nas emissões históricas de gases de efeito estufa, que são os principais responsáveis pelo aquecimento global.

O que podemos inferir:

Confiabilidade: O Climate Watch é uma fonte confiável de dados climáticos, utilizada por pesquisadores, governos e organizações internacionais. Para termos foco na nossa análise, iremos considerar apenas os dados que dizem respeito ao nosso país, Brasil.

Abrangência: O período de 30 anos permite analisar tendências de longo prazo nas emissões de GEE, o que é crucial para entender a evolução do problema e avaliar a eficácia das políticas climáticas.

2. Consumo de energia renovável (% do consumo total final de energia) - América Latina e Caribe⁽³⁾

- **Fonte:** Os dados foram compilados por um grupo de organizações internacionais, incluindo a Agência Internacional de Energia (IEA), a Agência Internacional de Energia Renovável (IRENA), a Divisão de Estatísticas das Nações Unidas (UNSD), o Banco Mundial e a Organização Mundial da Saúde (OMS).
- **Publicação:** Os dados foram publicados em 2023 no relatório "Rastreamento do ODS 7: O Relatório de Progresso Energético", produzido pelo Banco Mundial.
- **Licença:** O dataset é licenciado sob a Creative Commons Attribution, o que significa que podemos usá-lo, compartilhá-lo e modificá-lo, desde que atribua a autoria original.

O que podemos inferir:

Qualidade dos dados: A colaboração de várias organizações internacionais renomadas sugere que os dados são confiáveis e de alta qualidade.

Relevância: O dataset está diretamente relacionado ao Objetivo de Desenvolvimento Sustentável (ODS) 7, que busca garantir o acesso universal à energia elétrica e promover o uso de fontes de energia renováveis.

Para relacionar os dois datasets e desenvolver um projeto de machine learning eficaz, é fundamental compreender a fundo cada um deles. Quando falamos sobre previsões de emissões, utilizar dados históricos de consumo de energia renovável para prever futuras emissões de gases de efeito estufa. Pode-se utilizar modelos de séries temporais (ARIMA, LSTM, etc.) para identificar padrões e tendências nos dados de consumo de energia renovável e, em seguida, relacioná-los com as emissões de gases de efeito estufa.

4. REFERÊNCIAS BIBLIOGRÁFICAS

- (1) POYTING, Mark; PRAZERES, Leandro. Mudanças climáticas foram 'principal' fator para seca recorde na Amazônia, diz estudo: o que isso significa para o futuro da floresta?. [S. l.], 24 jan. 2024. Disponível em:
<https://www.bbc.com/portuguese/articles/c88nr0940j8o>. Acesso em: 11 set. 2024.
- (2) CO2 emissions (metric tons per capita) - Latin America & Caribbean. [S. l.], 14 fev. 2023. Disponível em:
<https://data.worldbank.org/indicator/EN.ATM.CO2E.PC?end=2020&locations=ZJ&start=1990&view=chart>. Acesso em: 11 set. 2024.
- (3) RENEWABLE energy consumption (% of total final energy consumption) - Latin America & Caribbean. [S. l.], 11 abr. 2023. Disponível em:
<https://data.worldbank.org/indicator/EG.FEC.RNEW.ZS?locations=ZJ>. Acesso em: 11 set. 2024.
- (4) ARAUJO, Arney Rayol Moura de; LOBATO, Tarcísio da Costa. Análise e previsão das emissões de CO2 provenientes dos Combustíveis do Setor de Transporte Rodoviário no Brasil 1990- 2020. [S. l.], 1 dez. 2020. Disponível em:
<http://repositorioinstitucional.uea.edu.br/bitstream/riuea/5280/1/Análise%20e%20previsão%20das%20emissões%20de%20CO2%20provenientes%20dos%20Combustíveis%20do%20Setor%20de%20Transporte%20Rodoviário%20no%20Brasil%201990-2020.pdf>. Acesso em: 30 set. 2024.
- (5) SOARES, Francisco. Aplicação de Machine Learning para Previsão das Emissões de CO2e Produzidas pela Produção de Energia no Estado do Amazonas: Um Estudo de Caso Usando R. [S. l.], 12 fev. 2024. Disponível em:
<https://medium.com/@fanciscosoares/aplicação-de-machine-learning-para-previsão-das-emissões-de-co2e-produzidas-pela-produção-de-1817cdd34729>. Acesso em: 30 set. 2024.

5. APÊNDICES

(1) Link do repositório do projeto:

<https://github.com/ViniSegatto/ProjetoAplicadoIV>