



Análise e previsão do consumo Energético em Indústria Usando Dados Históricos e Algoritmo Random Forest

Prof. Me. Elcio Rodrigues Aranha

Instituto Federal de Educação, Ciência e Tecnologia (IFSP), Cubatão, SP, Brasil.

Vitor Rabello Mendes

Instituto Federal de Educação, Ciência e Tecnologia (IFSP), Cubatão, SP, Brasil.

Resumo: Neste artigo, é abordado o problema de modelamento do consumo energético em uma indústria siderúrgica, utilizando dados históricos de eficiência energética. A solução proposta envolve o uso de um modelo de aprendizado de máquina, especificamente um regressor *Random Forest*, para prever o consumo de energia com base em diversas variáveis operacionais. Foi realizada uma análise dos dados para entender as correlações e distribuições das variáveis envolvidas. Em seguida, o modelo é treinado e testado, utilizando uma abordagem de divisão de dados, obtendo métricas de desempenho como MAE, MSE e R^2 que indicam a alta precisão do modelo. Conclui-se que a utilização do modelo *Random Forest* é uma alternativa eficaz na previsão do consumo energético e que a análise dos dados históricos pode auxiliar a indústria na identificação de padrões de consumo e na tomada de decisões para otimização energética.

Palavras-chave: *Random Forest*. Aprendizado de Máquina. Eficiência Energética.

Abstract: In this article, we address the issue of modeling energy consumption in a steel industry using historical data on energy efficiency. The proposed solution involves employing a machine learning model, specifically a Random Forest regressor, to predict energy consumption based on various operational variables. Data analysis was conducted to understand the correlations and distributions of the variables involved. Subsequently, the model was trained and tested using a data splitting approach, yielding performance metrics such as MAE, MSE, and R^2 that indicate high model accuracy. It will be demonstrated that the use of the Random Forest model is an effective alternative for predicting energy consumption, and that

analyzing historical data can assist the industry in identifying consumption patterns and making decisions for energy optimization.

Keywords: *Random Forest, Machine Learning, Energy Efficiency.*

INTRODUÇÃO

No contexto atual, a eficiência energética é uma das maiores preocupações em diversas indústrias, onde o consumo de energia tem um papel vital nas operações diárias. Tendo isso em mente, sabe-se que a demanda por recursos energéticos tem aumentado e a expectativa é que aumente ainda mais. Isso acaba gerando uma pressão nas empresas para otimizar seus processos industriais visando a reduzir os custos e minimizar o impacto ambiental (Tres *et al.*, 2021).

A busca por métodos mais eficazes de monitoramento e controle dos processos industriais é constante, e o uso de dados históricos para aprender padrões de consumo, identificar áreas de desperdício e implementar estratégias para melhorar a eficiência tem se destacado como uma ferramenta poderosa neste sentido (Camoto; Rebelatto, 2014).

Neste cenário, as árvores de decisão podem ser uma abordagem promissora para modelar e prever o comportamento de processos industriais com base em dados históricos de eficiência energética (Velloso; Hora, 2019). As árvores de decisão são modelos que replicam o processo de tomada de decisão humano, dividindo os dados em subconjuntos com base em características relevantes (Charbuty; Abdulazeez, 2021). O algoritmo *Random Forest* é uma técnica de aprendizado de máquina que combina várias árvores de decisão individuais para criar um modelo mais preciso (Breiman, 2001). Ao contrário de uma única árvore de decisão, que pode ser suscetível a *overfitting*, ou seja, ajustar-se excessivamente aos dados de treinamento e capturar ruídos e variações específicas, o *Random Forest* reduz esta instabilidade ao construir múltiplas árvores independentes e combinar suas previsões por meio de votação ou média (Nicola, 2021).

Este estudo propõe uma abordagem utilizando aprendizado de máquina para modelar e prever o consumo energético em uma indústria siderúrgica. A partir de dados históricos de eficiência energética, fornecidos pela *Korea Electric Power Corporation*, obtidos através da *Kaggle*, uma comunidade *on-line* voltada para cientistas de dados e entusiastas de aprendizado de máquina, conhecida por hospedar competições de ciência de dados, fornece conjuntos de dados e oferece

uma plataforma colaborativa para a prática e desenvolvimento de habilidades em análise de dados (Bojer; Meldgaard, 2021). A base de dados inclui medições detalhadas de diversas variáveis operacionais, como consumo de energia, potência reativa, emissão de CO₂, fator de potência, *status* semanal (final de semana ou dia útil), dia da semana e tipo de carga (leve, média, máxima). Estas medições foram coletadas em intervalos regulares ao longo de um ano, permitindo uma análise das condições operacionais e padrões de consumo energético. A volumosa quantidade de dados fornece uma base adequada para o desenvolvimento de modelos preditivos precisos e para a identificação de oportunidades de otimização energética na indústria siderúrgica.

O objetivo é desenvolver um modelo preciso que possa prever o consumo de energia com base em variáveis operacionais. Ao compreender estas relações, a indústria pode identificar padrões de consumo, antecipar necessidades energéticas e implementar estratégias de otimização.

Neste artigo, serão detalhados, na metodologia, a análise dos dados, o processo de treinamento e teste do modelo, e as métricas de avaliação utilizadas para validar a eficácia do modelo. A análise dos resultados obtidos e suas implicações para a gestão energética na indústria siderúrgica também são discutidos.

Em suma, este trabalho busca contribuir para o avanço do conhecimento no campo da eficiência energética industrial, fornecendo uma alternativa que possa ser aplicada na prática para otimização de processos e redução do consumo de energia.

METODOLOGIA

O objeto de estudo deste artigo é desenvolver um modelo de previsão do consumo energético de uma indústria siderúrgica, analisado através de dados históricos fornecidos pela *Korea Electric Power Corporation*. A base de dados contém medições de diversas variáveis operacionais coletadas em intervalos regulares.

A estratégia utilizada foi o desenvolvimento de um modelo preditivo baseado em aprendizado de máquina, especificamente um *Random Forest Regressor*. Esse modelo foi treinado para prever o consumo de energia (em kWh) utilizando diversas variáveis operacionais. No estudo, foram empregadas técnicas estatísticas (Meyer,

1982) e de aprendizado de máquina para analisar dados numéricos e gerar previsões sobre o consumo energético da indústria.

O pré-processamento dos dados envolveu várias etapas para garantir que os dados estivessem limpos e prontos para o modelo de aprendizado de máquina. Inicialmente, a coluna de data foi convertida para o formato *datetime*, facilitando a manipulação temporal dos dados. Em seguida, valores ausentes e duplicados foram tratados para evitar distorções nas análises subsequentes. Variáveis categóricas foram codificadas utilizando o *LabelEncoder*, transformando dados categóricos em numéricos. As variáveis contínuas foram escalonadas para garantir que todas tivessem a mesma influência no modelo, utilizando o *StandardScaler*. Finalmente, os dados foram divididos em conjuntos de treinamento e teste para validar a performance do modelo preditivo.

Para avaliar a precisão e a eficácia do modelo de previsão de consumo energético, foram utilizadas três métricas de desempenho comuns: *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE) e *Coefficient of Determination* (R^2). Cada uma dessas métricas fornece uma perspectiva diferente sobre o desempenho do modelo.

O MAE é a média dos erros absolutos entre os valores preditos pelo modelo e os valores reais. Ele fornece uma medida intuitiva da precisão do modelo em termos das unidades da variável alvo. Um MAE menor indica um modelo mais preciso, pois os erros são, em média, menores (Chicco; Warrens; Jurman, 2021).

O MSE é a média dos quadrados dos erros entre os valores preditos e os valores reais. A elevação ao quadrado amplifica os erros maiores, tornando o MSE mais sensível a grandes desvios. Um MSE menor indica um modelo com menos desvios significativos entre os valores reais e preditos (Chicco; Warrens; Jurman, 2021).

O R^2 representa a proporção da variância explicada pelo modelo, sempre variando entre 0 e 1, onde 1 indica que o modelo explica toda a variância dos dados e 0 indica que o modelo não explica nenhuma variância. É independente da escala de Y, isso torna o R^2 uma medida mais intuitiva e comparável do desempenho do modelo. Um R^2 próximo de 1 indica que o modelo tem um bom ajuste e explica bem a variância dos dados. Um valor próximo de 0 indica um ajuste ruim (James *et al.*, 2013).

A amostra consiste em um conjunto de dados contendo 35.040 linhas por 11 colunas, cada uma representando medições de diversas variáveis operacionais em um determinado momento. Os dados foram coletados ao longo de um período, abrangendo diferentes condições operacionais da indústria.

Para o processamento e análise dos dados, foram utilizadas as seguintes bibliotecas de *Python*:

Pandas: para manipulação e análise de dados, fornece estruturas de dados flexíveis e expressivas, como *DataFrames*, que facilitam a limpeza, transformação e análise dos dados. É essencial para a leitura de arquivos CSV, tratamento de dados ausentes e execução de operações complexas de manipulação de dados.

NumPy: Usada para suporte a *arrays* e matrizes multidimensionais, além de fornecer uma coleção abrangente de funções matemáticas para operações rápidas e eficientes. É a base para muitas outras bibliotecas de ciência de dados, oferecendo funcionalidade fundamental para cálculos numéricos.

Matplotlib: Biblioteca de visualização de dados que permite a criação de gráficos estáticos, animados e interativos em *Python*. É extremamente útil para gerar gráficos de linha, dispersão, histogramas e outros tipos de visualizações para explorar e entender os dados.

Seaborn: Baseada no *Matplotlib*, fornece uma interface de alto nível para criar visualizações estatísticas atraentes e informativas. Facilita a geração de gráficos complexos com menos código e inclui suporte para temas estéticos que melhoram a aparência dos gráficos.

Plotly: Biblioteca de visualização interativa que permite a criação de gráficos dinâmicos e interativos, incluindo gráficos de linha, dispersão, mapas e muito mais. É especialmente útil para apresentações e *dashboards* interativos que permitem uma exploração mais profunda dos dados.

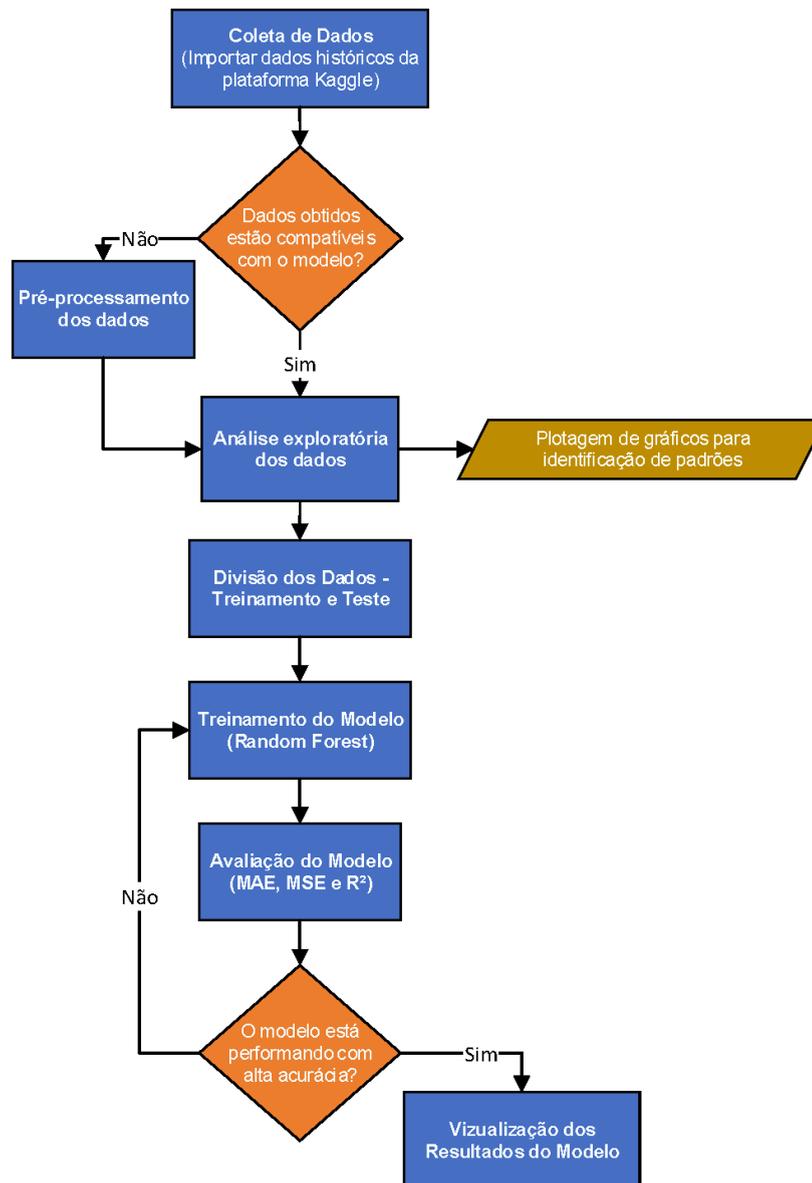
Scikit-learn: Biblioteca de aprendizado de máquina que fornece ferramentas simples e eficientes para análise de dados e mineração de dados. Inclui uma ampla variedade de algoritmos de aprendizado supervisionado e não supervisionado, ferramentas de validação cruzada, seleção de características e mais. É fundamental para o desenvolvimento, treinamento e avaliação de modelos de aprendizado de máquina.

A programação foi realizada no *Google Colab*, que fornece um ambiente de desenvolvimento interativo e baseado em nuvem.

DESENVOLVIMENTO

Antes de apresentar o desenvolvimento, é importante entender a sequência de etapas seguidas para desenvolver e avaliar o modelo de previsão de consumo energético. O fluxograma na Fig. 1 ilustra essas etapas de maneira clara e organizada.

Figura 1 – Fluxograma do desenvolvimento



Fonte: Elaborado pelos autores (2024).

Os dados coletados forneciam as seguintes variáveis:

- *Date*: Data referente a extração dos dados
- *Usage_kWh*: Consumo de energia em kWh (variável dependente)
- *Lagging Current Reactive Power_kVArh*: Potência reativa atrasada em kVArh
- *Leading Current Reactive Power_kVArh*: Potência reativa adiantada em kVArh

- CO2 (tCO2): Níveis de CO2 em toneladas
- *Lagging Current Power Factor*: Fator de potência atrasado
- *Leading Current Power Factor*: Fator de potência adiantado
- NSM: Número de segundos desde a meia-noite
- *WeekStatus*: Status da semana (fim de semana ou dia de semana)
- *Day_of_week*: Dia da semana
- *Load_Type*: Tipo de carga (leve, média, máxima)

O desenvolvimento do modelo iniciou-se com a importação das bibliotecas necessárias e o carregamento da base de dados.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import plotly.express as px
import seaborn as sns
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.feature_selection import RFE
from sklearn.tree import DecisionTreeRegressor
from sklearn.metrics import *

#Carregar os Dados
df = pd.read_csv('Steel_industry_data.csv')
```

Após a visualização dos dados na plataforma *Google Colab*, identificou-se a necessidade de tratamento da variável "Date" para possibilitar uma melhor análise dos dados fornecidos.

```
# Converter a coluna de data para o formato datetime
df['date'] = pd.to_datetime(df['date'], format = "%d/%m/%Y
%H:%M")
```

Foi verificada a presença de valores duplicados ou nulos nos dados, além de realizado o tratamento das variáveis categóricas.

```
# Verificar valores duplicados e nulos
df.duplicated().sum()
df.isna().sum()

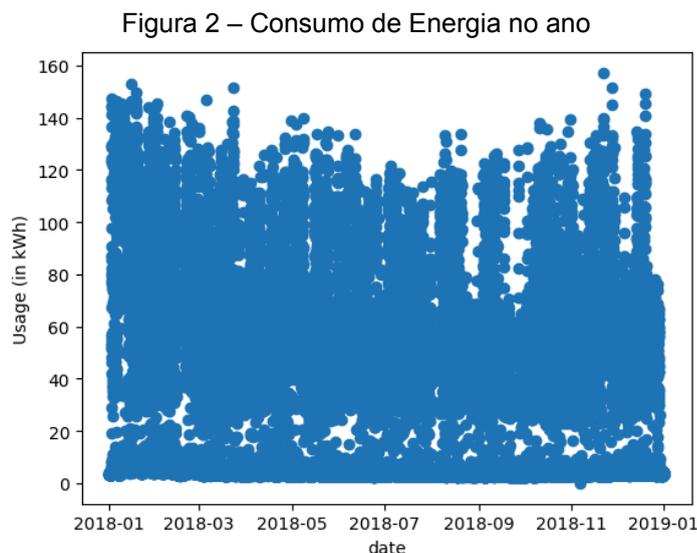
# Codificar variáveis categóricas
le = LabelEncoder()
for i in df.columns:
    if df[i].dtypes == 'object':
        df[i] = le.fit_transform(df[i])
```

As variáveis categóricas 'WeekStatus', 'Day_of_week' e 'Load_Type' foram transformadas em valores numéricos utilizando o *LabelEncoder*, possibilitando seu uso no modelo de aprendizado de máquina.

Foi realizada uma análise para entender melhor as relações entre as variáveis fornecidas.

```
# Gráficos de dispersão para visualizar a relação entre cada
variável e Usage_kWh
for i in df.columns:
    if i != 'Usage_kWh':
        plt.scatter(x = df[i], y = df['Usage_kWh'])
        plt.xlabel(i)
        plt.ylabel('Usage (in kWh)')
        plt.show()
```

Foram gerados os gráficos, os quais alguns estão sendo representados nas Fig. 2, 3, 4, 5 e 6 com o objetivo de identificar possíveis padrões e relações entre as variáveis em relação ao consumo de energia em kWh.

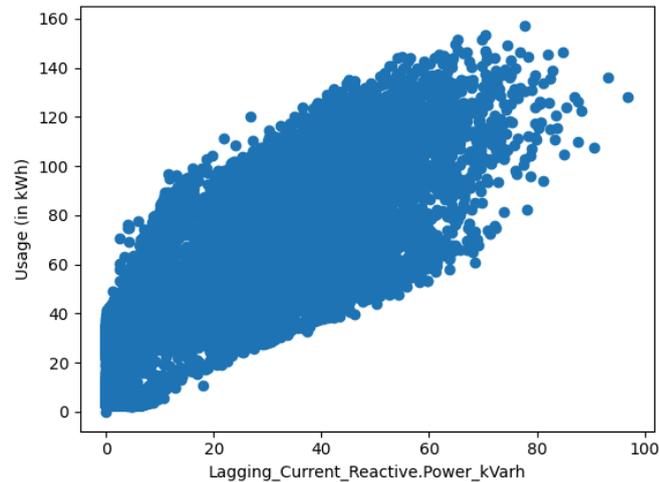


Fonte: Elaborado pelos autores (2024).

Cada ponto no gráfico representa uma observação individual da base de dados. A Fig. 2 mostra a distribuição do consumo de energia (em kWh) ao longo do ano. Observa-se uma ampla variação nos valores de consumo, com alguns picos notáveis, especialmente em períodos específicos do ano. Esta variação pode ser influenciada por fatores sazonais, condições operacionais da indústria e eventos excepcionais que podem ter aumentado o consumo de energia.

A Fig. 3 ilustra a relação entre o consumo de energia e a potência reativa atrasada. Note-se uma correlação positiva clara, conforme a potência reativa atrasada aumenta, o consumo de energia também tende a aumentar. Isso sugere que a potência reativa atrasada é um dos fatores importantes a serem considerados ao prever o consumo de energia na indústria.

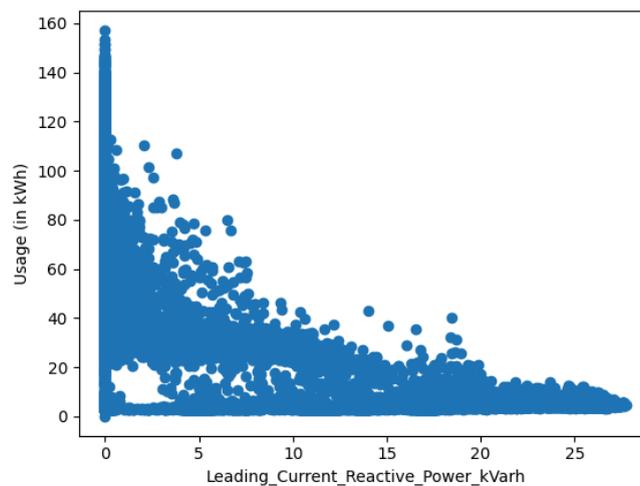
Figura 3 – Consumo de Energia com potência reativa atrasada



Fonte: Elaborado pelos autores (2024).

A Fig. 4 apresenta a relação entre o consumo de energia e a potência reativa adiantada. Diferentemente da potência reativa atrasada, há uma correlação negativa, quando a potência reativa adiantada é baixa, o consumo de energia tende a ser maior, mas conforme a potência reativa adiantada aumenta, o consumo de energia diminui.

Figura 4 – Consumo de Energia com potência reativa adiantada

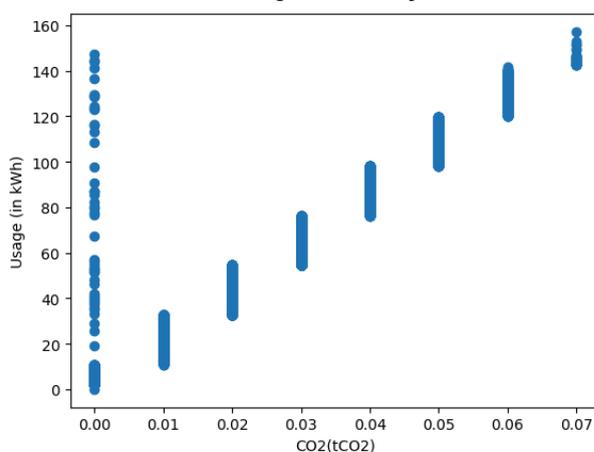


Fonte: Elaborado pelos autores (2024).

Esta inversão na tendência pode ser devido a diferentes comportamentos de carga e gerenciamento de energia reativa na operação industrial.

Na Fig. 5 mostra-se uma clara tendência crescente, indicando que, à medida que a emissão de CO₂ aumenta, o consumo de energia também aumenta. Esta correlação positiva sugere que atividades com maior consumo energético são diretamente responsáveis por maiores emissões de CO₂. Isso pode ser devido ao uso de combustíveis fósseis ou outros processos industriais intensivos em energia que geram CO₂. No modelo desenvolvido, a variável (emissão de CO₂) pode ser um preditor significativo para o consumo de energia. A correlação visível ajuda a justificar a inclusão desta variável no modelo de previsão.

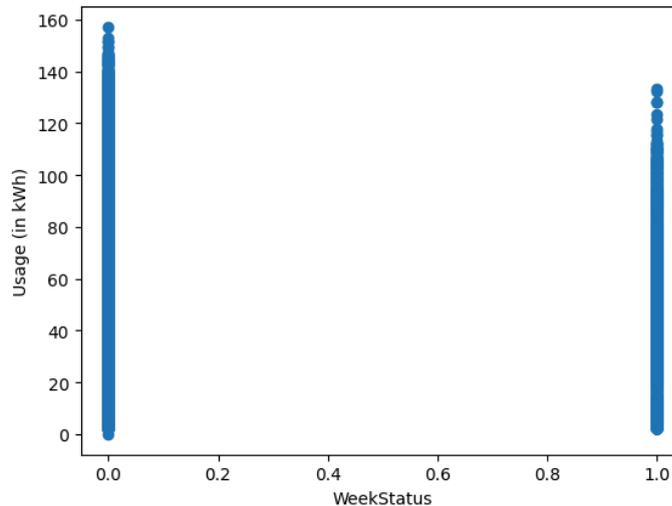
Figura 5 – Consumo de energia em relação a emissão de CO₂



Fonte: Elaborado pelos autores (2024).

A Fig. 6 mostra o consumo de energia em diferentes momentos da semana, diferenciando entre dias de semana (0) e fins de semana (1). A visualização dos dados indica que o consumo de energia é similar nos dias de semana e fins de semana, com uma concentração densa de pontos em ambos os extremos. Isso sugere que o padrão de consumo de energia não varia significativamente entre dias de semana e fins de semana, mantendo níveis altos e consistentes de uso de energia em ambos os períodos.

Figura 6 – Consumo de energia pelo status da semana
(0 = Dia de semana; 1 = Final de Semana)



Fonte: Elaborado pelos autores (2024).

A análise dos gráficos plotados revelou alguns padrões essenciais para um treinamento eficaz do modelo.

Compreendendo como a indústria do exemplo apresentado funciona em termos de consumo de energia, foi possível trabalhar na predição de possíveis mudanças na estrutura. Portanto, os dados foram divididos em conjuntos de treino e teste, e a coluna de data foi removida para o treinamento do modelo, embora ainda seja utilizada ao longo do algoritmo.

```
#Definir variável com dados da coluna date
df_date = df['date']

# Remover a coluna de data para treinamento
df = df.drop(['date'], axis = 1)

# Separar dados em características e rótulo
x = df.drop(['Usage_kWh'], axis = 1)
y = df['Usage_kWh']

# Dividir dados em treinamento e teste
x_train,x_test,y_train,y_test =
train_test_split(x,y,train_size=0.75, bsolu_state=10)
```

Após a coluna “date” ser removida e os dados serem divididos em variáveis independentes (X) e dependentes (y), estes também foram divididos em conjuntos de treino e teste.

Na etapa seguinte, utilizou-se o algoritmo *Random Forest Regressor* para treinar o modelo.

```
From sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error,
mean_squared_error
# Treinamento do modelo
model = RandomForestRegressor(n_estimators=100)
model.fit(x_train, y_train)
```

O modelo *Random Forest Regressor* é instanciado e treinado utilizando os dados de treino.

Logo em seguida, o modelo foi validado utilizando métricas de desempenho.

```
# Predição
y_pred = model.predict(x_test)

# Comparar as previsões com os valores reais
mae = mean_absolute_error(y_test, y_pred)
mse = mean_squared_error(y_test, y_pred)
r2 = model.score(x_test, y_test)

print(f'MAE: {mae}')
print(f'MSE: {mse}')
print(f'R²: {r2}')
```

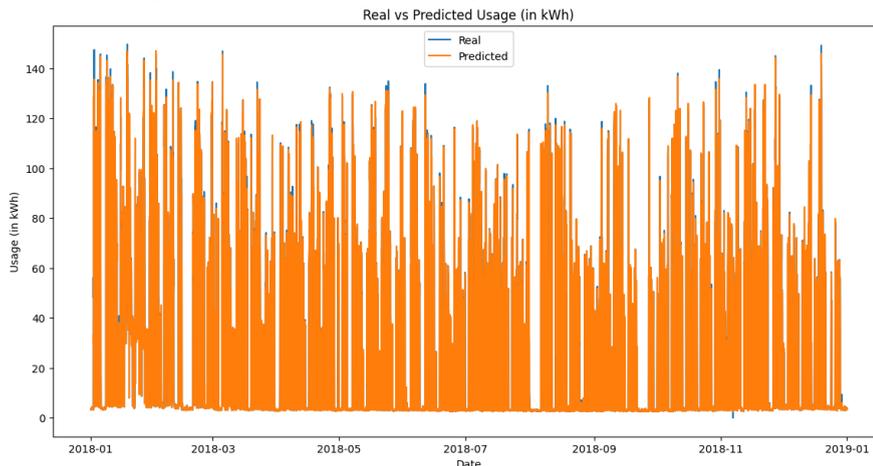
A predição é realizada no conjunto de teste e as métricas MAE, MSE e R^2 são calculadas para avaliar o desempenho do modelo, os valores das métricas foram:

- MAE: 0.3363035958904112
- MSE: 1.140205025886988
- R^2 : 0.9990037192162694

Estes resultados indicam que o modelo tem uma alta precisão e um excelente ajuste aos dados, sendo eficaz na previsão do consumo energético. O MAE e o MSE baixos indicam que os erros de predição são pequenos, e o R^2 próximo de 1 mostra que o modelo explica quase toda a variância dos dados de consumo energético.

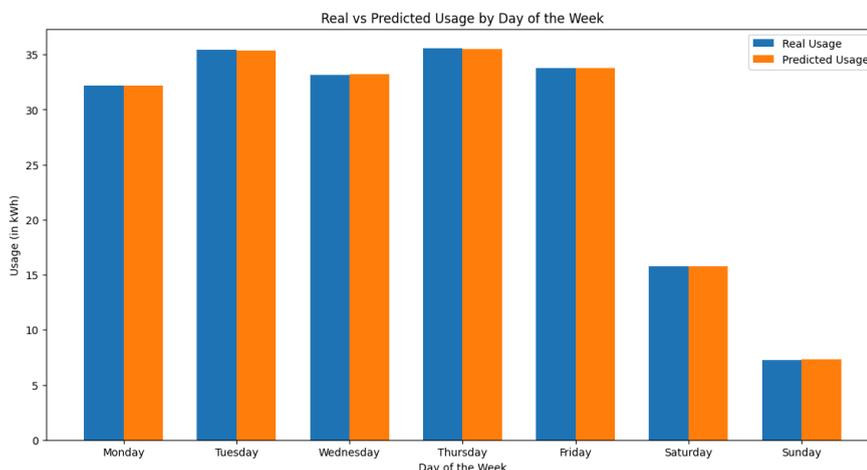
Para visualização dos resultados, foram criados gráficos comparativos entre os valores reais e os valores preditos pelo modelo, facilitando a compreensão dos dados. A proximidade entre as linhas azul (real) e laranja (predito) na Fig. 7 indica quão bem o modelo *Random Forest Regressor* conseguiu prever o consumo de energia. Quanto mais próximas estiverem as linhas, mais preciso é o modelo. Note-se que as linhas parecem estar bastante sobrepostas, sugerindo que o modelo tem um bom desempenho na previsão do consumo de energia.

Figura 7 – Consumo de energia no ano Real x Predito



Fonte: Elaborado pelos autores (2024).

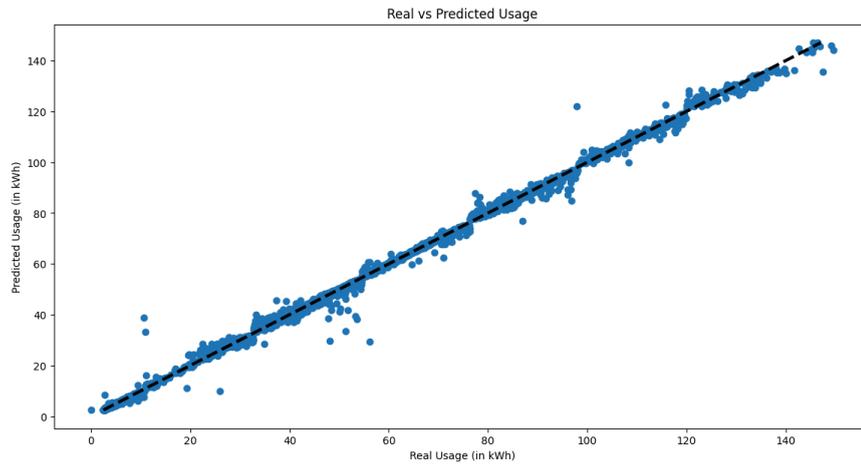
Figura 8 – Consumo de energia por dia da semana Real x Predito



Fonte: Elaborado pelos autores (2024).

A Fig. 9 apresenta um gráfico de dispersão que mostra uma forte correlação linear entre os valores reais e preditos de consumo de energia. A linha de tendência está próxima da linha de identidade, o que sugere que o modelo geralmente faz boas previsões. No entanto, há alguns *outliers*, principalmente em valores de consumo mais altos, onde o modelo tende a subestimar o consumo real. Isso sugere que o modelo pode precisar de ajustes para melhorar a previsão em cenários de consumo elevado.

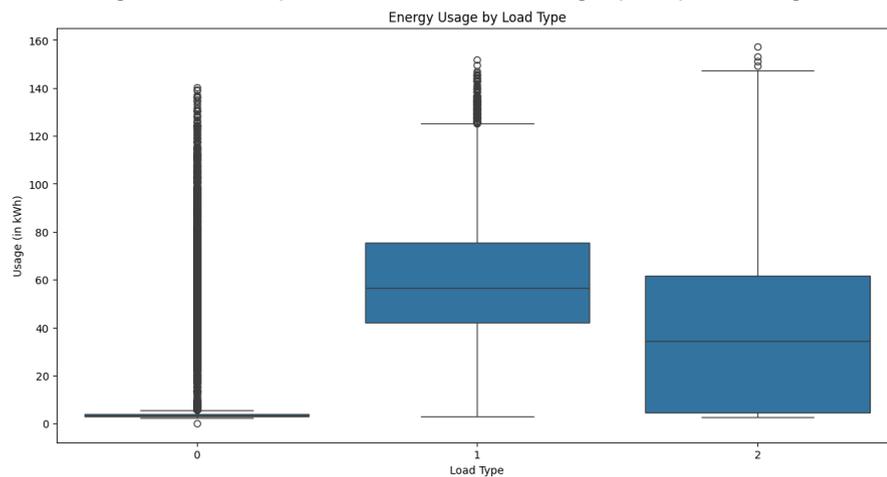
Figura 9 – Gráfico de dispersão do consumo de energia Real x Predito



Fonte: Elaborado pelos autores (2024).

Um *boxplot*, representado na Fig. 10, compara o consumo de energia por tipo de carga. Cargas leves (0) apresentam uma distribuição com muitos *outliers* e uma mediana baixa, indicando que a maioria dos consumos são baixos, mas há algumas ocorrências de consumo muito alto. Cargas médias (1) têm uma mediana mais alta e uma maior variação de consumo comparada às cargas leves. Cargas máximas (2) mostram a maior variação de consumo e apresentam muitos *outliers*. A análise por tipo de carga revela que diferentes perfis de carga têm variações distintas de consumo, indicando que o modelo pode precisar considerar mais detalhadamente essas diferenças para aprimorar a precisão das previsões.

Figura 10 – *Boxplot* do consumo de energia por tipo de carga



Fonte: Elaborado pelos autores (2024).

CONCLUSÃO

Neste estudo, foi abordado o desafio de modelar o consumo energético em uma indústria siderúrgica utilizando dados históricos de eficiência energética. Foi implementado um modelo de aprendizado de máquina, mais precisamente um *regressor Random Forest*, para prever o consumo de energia com base em diversas variáveis operacionais. Os resultados indicam que o modelo apresenta alta precisão e eficácia na previsão do consumo energético, conforme evidenciado pelas métricas de desempenho obtidas: MAE de 0.3363, MSE de 1.1402 e R^2 de 0.9990.

Os gráficos comparando os valores reais e preditos demonstram que o modelo é capaz de capturar com precisão as variações diárias no consumo de energia. A análise visual reforça a confiança na capacidade do modelo de fazer previsões precisas, o que é necessário para validar o seu uso na prática.

A aplicação deste modelo de previsão pode proporcionar diversos benefícios para empresas que buscam melhorar sua eficiência energética. Com um modelo de alta precisão, como o desenvolvido neste estudo, as empresas podem:

- Prever o consumo futuro: antecipar o consumo energético com precisão permite que as empresas planejem melhor suas operações, ajustando a produção e os horários de funcionamento para otimizar o uso de energia.
- Identificar padrões de consumo: a análise dos dados históricos e as previsões geradas pelo modelo ajudam a identificar padrões de consumo energético. Isso pode revelar oportunidades para implementar medidas de eficiência energética, como a manutenção preventiva de equipamentos ou a substituição de tecnologias ineficientes.
- Reduzir custos operacionais: ao prever com precisão os picos de consumo e ajustar as operações para suavizar estas demandas, as empresas podem reduzir os custos associados ao consumo de energia, evitando sobretaxas e tarifas elevadas.

Este estudo evidencia o potencial de tais modelos em auxiliar empresas a otimizar seu uso de energia, reduzir custos e melhorar a eficiência operacional, contribuindo, assim, para a sustentabilidade e a competitividade no setor industrial.

REFERÊNCIAS

BOJER, C. S.; MELDGAARD, J. P. Kaggle forecasting competitions: An overlooked learning opportunity. **International Journal of Forecasting**, v. 37, n. 2, p. 587-603, 2021.

BREIMAN, Leo. Random forests. **Machine learning**, v. 45, p. 5-32, 2001.

CAMIOTO, F. de C.; REBELATTO, D. A. do N. Análise da contribuição ambiental por meio da alteração da matriz energética do setor brasileiro de ferro-gusa e aço. **Gestão & Produção**, v. 21, p. 732-744, 2014.

CHARBUTY, B.; ABDULAZEEZ, A. Classification based on decision tree algorithm for machine learning. **Journal of Applied Science and Technology Trends**, v. 2, n. 01, p. 20-28, 2021.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. **PeerJ computer science**, v. 7, p. e623, 2021.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning With Applications in R**. New York: Springer, 2013.

MEYER, P. L. **Probabilidade: aplicações à estatística**. Livros Técnicos e Científicos, 1982.

NICOLA, M. J. **Adoção de random forest e regressão linear para previsão de falhas em equipamentos agrícolas**. 2021. 92f. Dissertação (Mestrado - Programa de Pós-Graduação em Mestrado Profissional em Matemática, Estatística e Computação Aplicadas à Indústria) - Universidade de São Paulo - São Carlos, 2021.

TRES, N; ZANIN, A.; KRUGER, S. D.; MAGRO, C. B. D. Sustainability practices adopted by industrial companies. **Revista de Administração da UFSM**, v. 14, n. spe, p. 1140-1159, 2021.

VELLOSO, H. M.; HORA, H. R. M. da. Classificação de falhas de um centro de usinagem: um estudo de caso utilizando árvore de decisão. *In*: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA, 19., 2019, Rio de Janeiro, RJ. **Anais [...]**. Rio de Janeiro: Centro de Análises de Sistemas Navais, 2019.