

OTIMIZADOR DE CONSUMO ENERGÉTICO - ANÁLISE PREDITIVA 2024

Resumo Executivo

- **Resultado Chave:** O modelo preditivo (V4: XGBoost-Only) alcançou um **Erro Percentual Médio (MAPE)** de **6.70%** no agregado global, com todas as regiões apresentando erro de um dígito.
- **Descoberta Principal:** O consumo é impulsionado por **1) População** (correlação de 0.96) e **2) Clima Regional**, que possui um impacto não-linear (ex: frio aumenta o consumo no Sul, calor aumenta no Sudeste).
- **Descoberta NLP:** Um pipeline de NLP foi desenvolvido para monitorar notícias em tempo real. Ele filtrou 100 artigos e identificou **6 eventos relevantes** (ex: alertas do Inmet), fornecendo contexto qualitativo para os picos de demanda.
- **Recomendação:** Implementar o modelo XGBoost treinado por região para otimizar o planejamento de despacho de carga diário, usando os alertas de NLP para validar picos de erro do modelo.

Metodologia

- **Dados Quantitativos (ML):** Três fontes de dados foram combinadas:
 1. consumo_historico_por_regiao.csv (Consumo)
 2. crescimento_populacional_Regioes_2020_2024.csv (População)
 3. medias_temperatura_umidade_2024.csv (Clima)
- **Dados Qualitativos (NLP):** Coleta em tempo real via API de Notícias (NewsAPI), filtrada por palavras-chave (apagão, onda de calor, ANEEL, etc.).
- **Algoritmo Preditivo (ML): XGBoost Regressor (Modelo V4).** Modelos híbridos (Prophet+XGBoost) foram testados (V1-V3) e descartados devido à instabilidade (MAPE > 1000%).
- **Algoritmo de Contexto (NLP): spaCy (pt_core_news_lg)** para Classificação de Tema e Extração de Entidades (NER).

Performance Detalhada do Modelo (ML)

A tabela abaixo detalha o erro (MAPE) do modelo final (V4) no período de teste:

| Região | Erro Médio (MAPE) | Erro Médio (MAE) | Insight Chave do Modelo |
|----------|-------------------|------------------|-----------------------------|
| Nordeste | 5.04 % | 410.73 MWm | Previsão altamente precisa. |

| | | | |
|-------------------|---------------|-------------------|---|
| Norte | 5.54 % | 219.38 MWm | Performance excelente. |
| Sudeste/CO | 7.38 % | 1649.37 MWm | Bom resultado, capturou o "efeito ar-condicionado". |
| Sul | 8.85 % | 605.41 MWm | Bom resultado, capturou o "efeito aquecedor". |
| GLOBAL | 6.70 % | 721.22 MWm | Performance validada. |

Análise Qualitativa (NLP)

O pipeline de NLP foi executado (src/data_collection.py + nlp_analysis.ipynb) para fornecer contexto aos dados.

- **Resultados da Coleta:** 100 notícias coletadas da API.
- **Resultados da Filtragem (NLP):** 94% dos artigos foram classificados como "Outros" (ruído, ex: futebol) e removidos.
- **Indicadores Relevantes Encontrados:** 6 artigos foram validados como relevantes, fornecendo contexto para anomalias na rede.

Amostra de Indicadores Relevantes Identificados (04-06/Novembro):

| Data | Manchete (Resumida) | Tema (NLP) | Entidades (NLP) |
|-------|--|--------------------------|---|
| 06/11 | Inmet: 9 estados têm alerta laranja... | Evento Climático Extremo | (Inmet, LOC), (Instituto Nacional de Meteorologia, ORG) |
| 06/11 | Lucro líquido da Alupar... | Regulatório/Empresarial | [] |

| | | | |
|-------|--------------------------------------|--------------------------|--|
| 06/11 | Inmet emite dois alertas laranjas... | Evento Climático Extremo | (Inmet, LOC), (Instituto Nacional de Meteorologia, ORG) |
|-------|--------------------------------------|--------------------------|--|

Fatores Chave Identificados (Drivers de Consumo)

1. **População (Peso: Alto - Macro):** Correlação de **0.96** com o consumo total.
 2. **Clima (Peso: Alto - Regional):** Driver não-linear (frio no Sul, calor no Sudeste).
 3. **Sazonalidade (Peso: Médio - Diário):** dia_semana e e_feriado (fins de semana/feriados).
 4. **Eventos Externos (Peso: Pontual - NLP):** Alertas de clima (Inmet) e problemas operacionais (apagão) que explicam picos de erro do modelo.
-

Recomendações Estratégicas (Otimização)

- **Curto Prazo (Opex):** Integrar o dashboard.py ao Centro de Operação. Usar a previsão diária (ML) para otimizar o despacho de carga e reduzir custos com usinas térmicas . Cruzar picos de erro (>10%) com os Indicadores Relevantes (NLP) para entender a causa (ex: "O modelo errou porque houve um alerta de calor que ele não previa").
- **Médio Prazo (Capex):** Usar os drivers do modelo (População e Clima) para simular cenários de expansão de infraestrutura e identificar onde a demanda crescerá mais nos próximos 5-10 anos.

Confiabilidade da Análise • Erro do Modelo (MAPE Global): 6.70% • Margem de Erro (MAE Global): ± 721.22 MWm

Confiabilidade da Análise

- Erro do Modelo (MAPE Global): 6.70%
- Margem de Erro (MAE Global): ± 721.22 MWm