

PREVISÃO DE ELDT

Grupo Stratus Data

Gabriel Scatena
Yan Teixeira

Etapas do Projeto

1

Montar a base de
treino

Maior preocupação
foi em ter uma base
sem qualquer data
leakage

2

Criar novos
atributos

Features
disponíveis não
tinham sinal

3

Diminuir o ruído

Diversos casos de
registros extremos
de ELDT

4

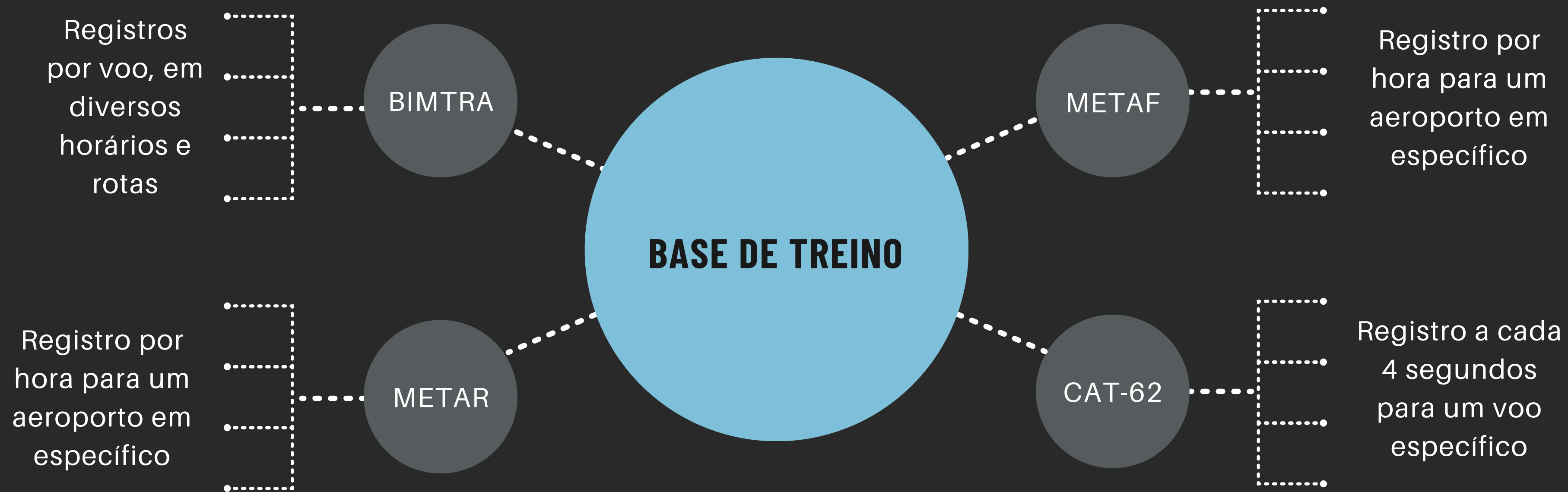
Escolher o modelo

Foi escolhido o
modelo que melhor
se adequava com a
base final de treino

5

Treino

Por fim, treino e
inferência do
modelo



BASE DE TREINO

BIMTRA

METAF

METAR

CAT-62

Horário
Referência da
decolagem

Aeroporto de
Destino

hora_ref + 1
Aeroporto de
Destino

Número de
aeronaves no
espaço aéreo
no momento
hora_ref

BASE DE TREINO

BIMTRA, Esperas, TC-prev

- Origem
- Destino
- Hora
- Espera
- Troca de Cabeceira

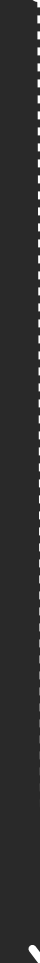
METAR, METAF

- Direção do vento
- Velocidade do vento
- Dew Point
- Visibilidade
- Temperatura

CAT-62

- Número de aeronaves em voo no instante hora_ref

Decolagem = Captura dos dados



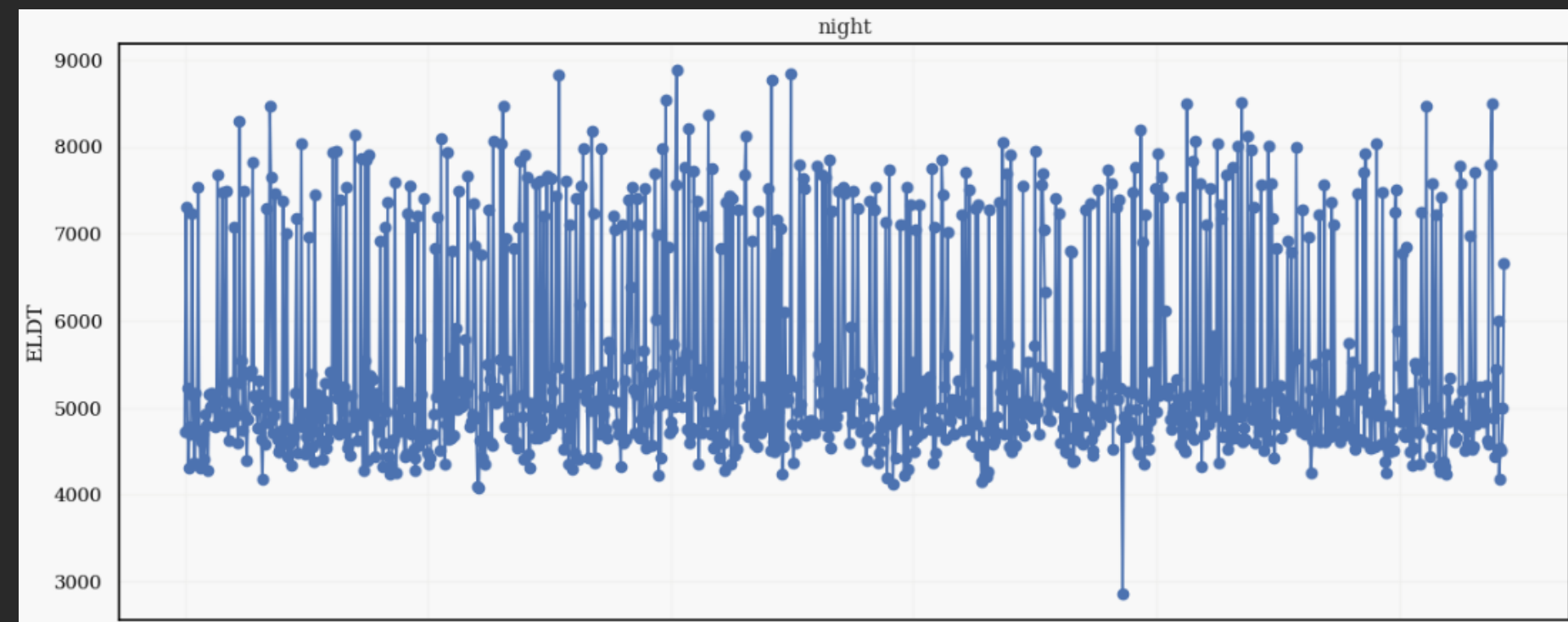
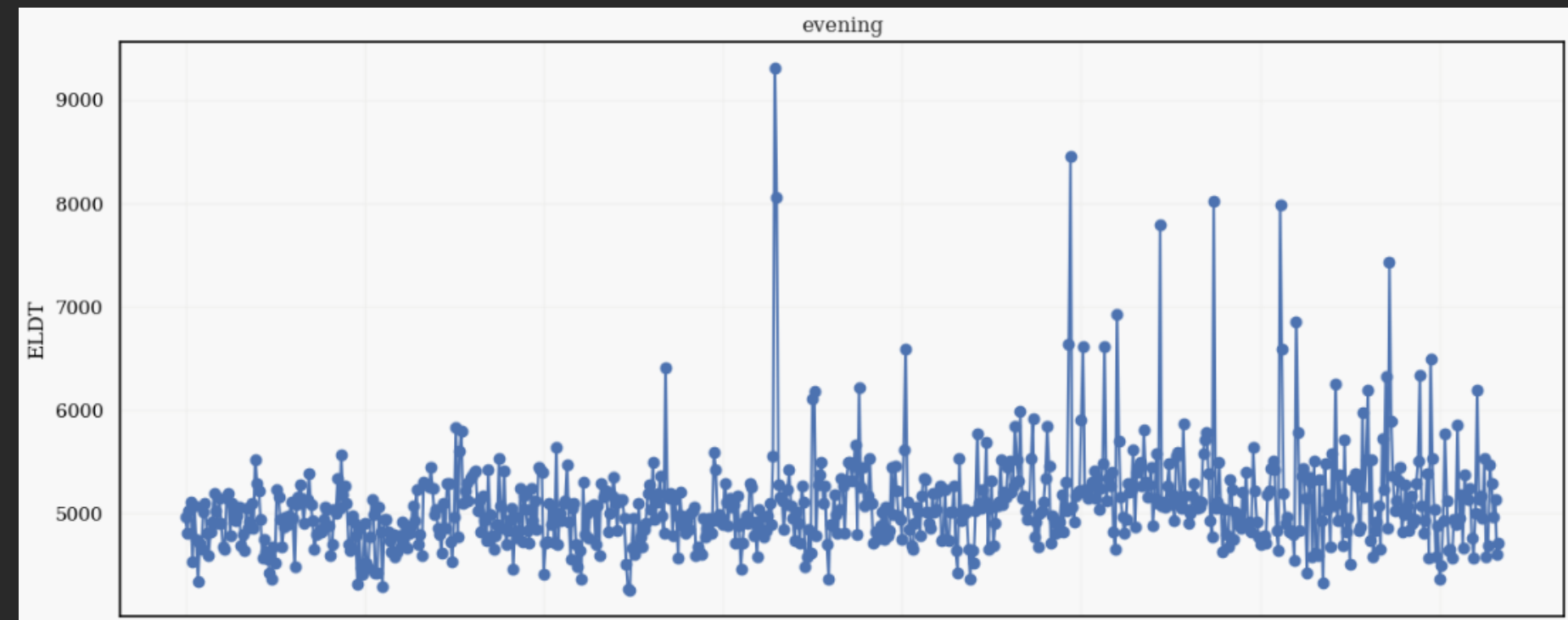
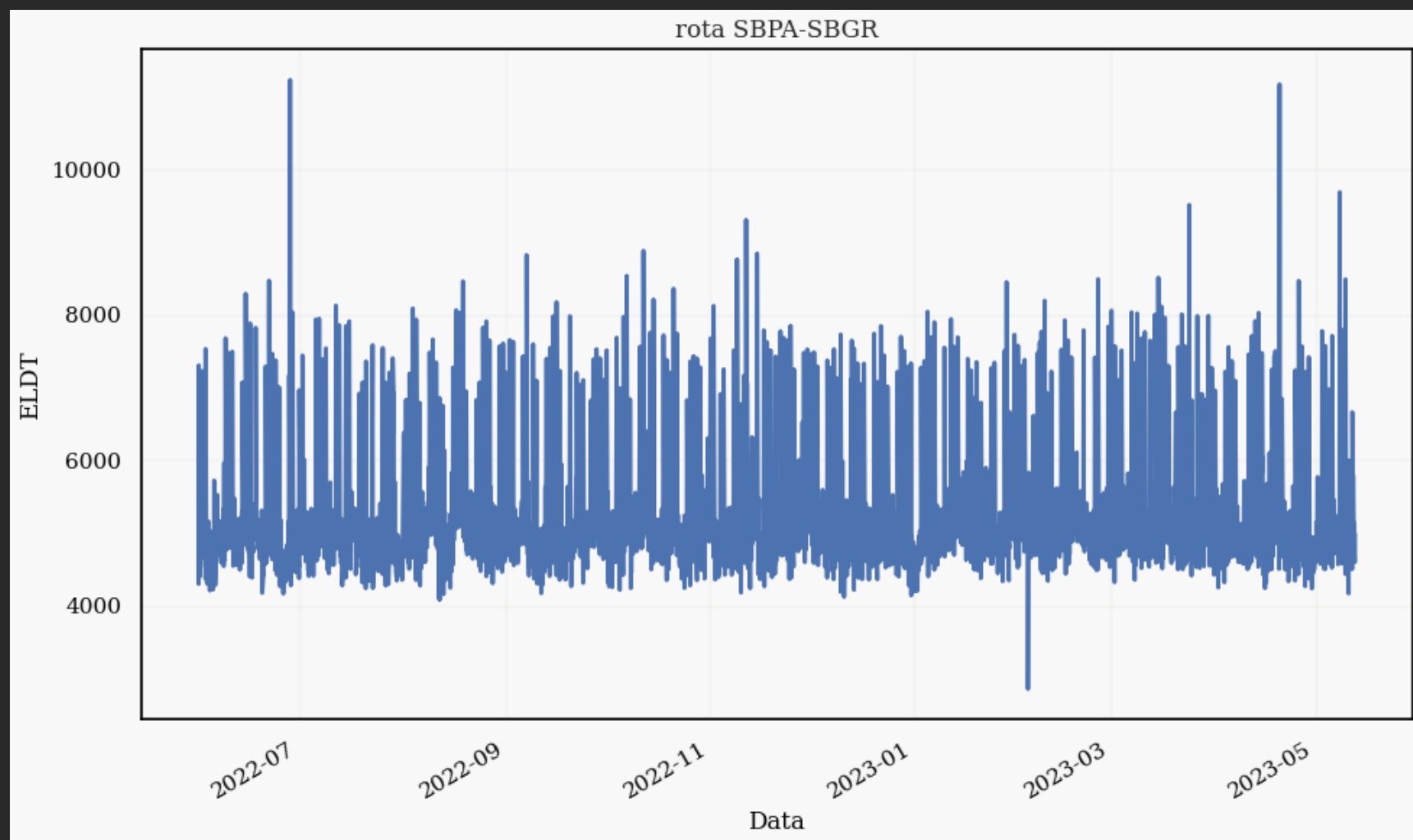
Após a decolagem não há mais informações disponíveis

Feature Engineering

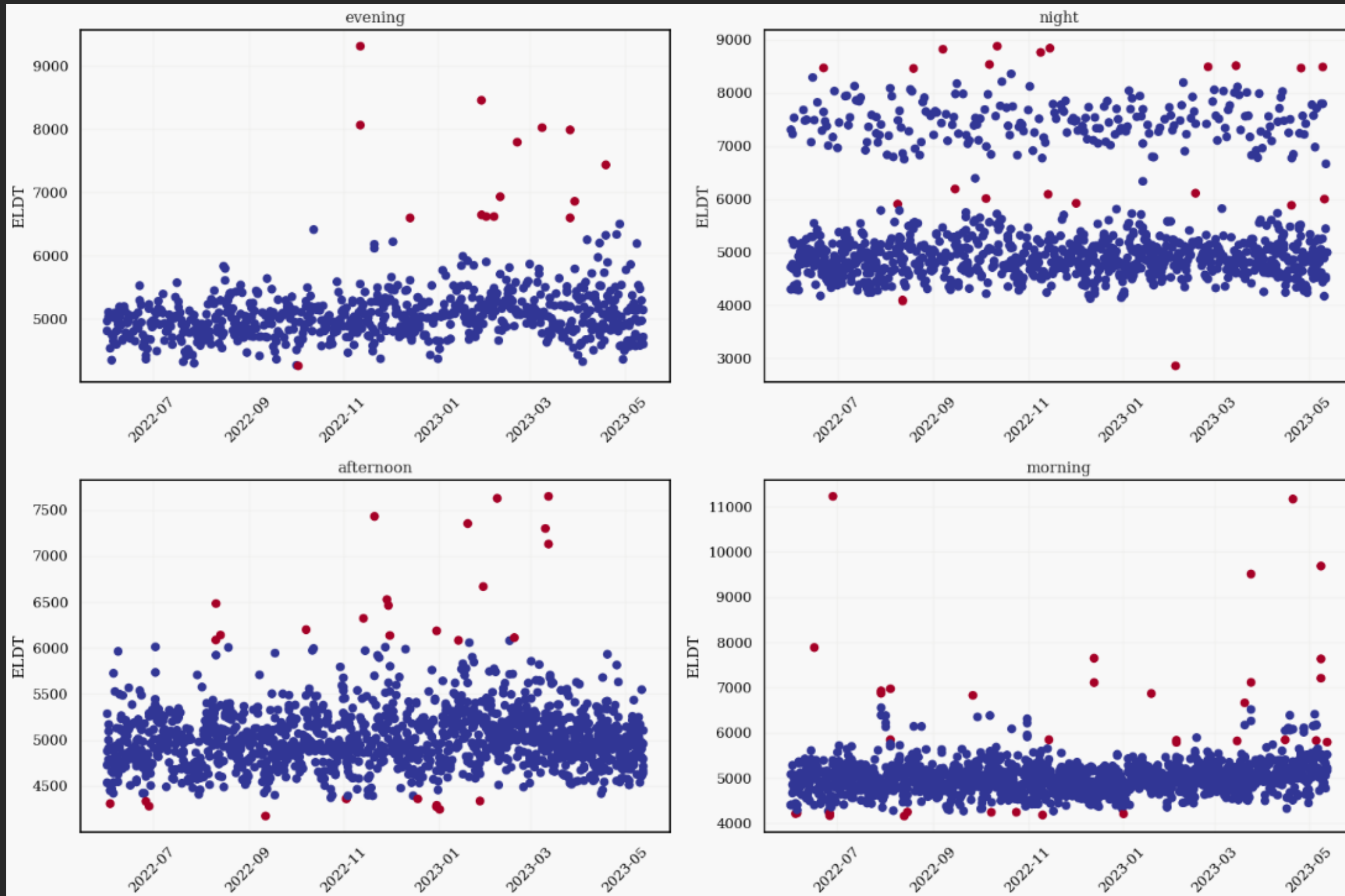


- Latitude da Origem
- Longitude da Origem
- Direção em graus do vetor Origem -> Destino
- Rota (categórica) = Origem + Destino
- Latitude do Destino
- Longitude do Destino

Pontos Extremos



Local Outlier Factor

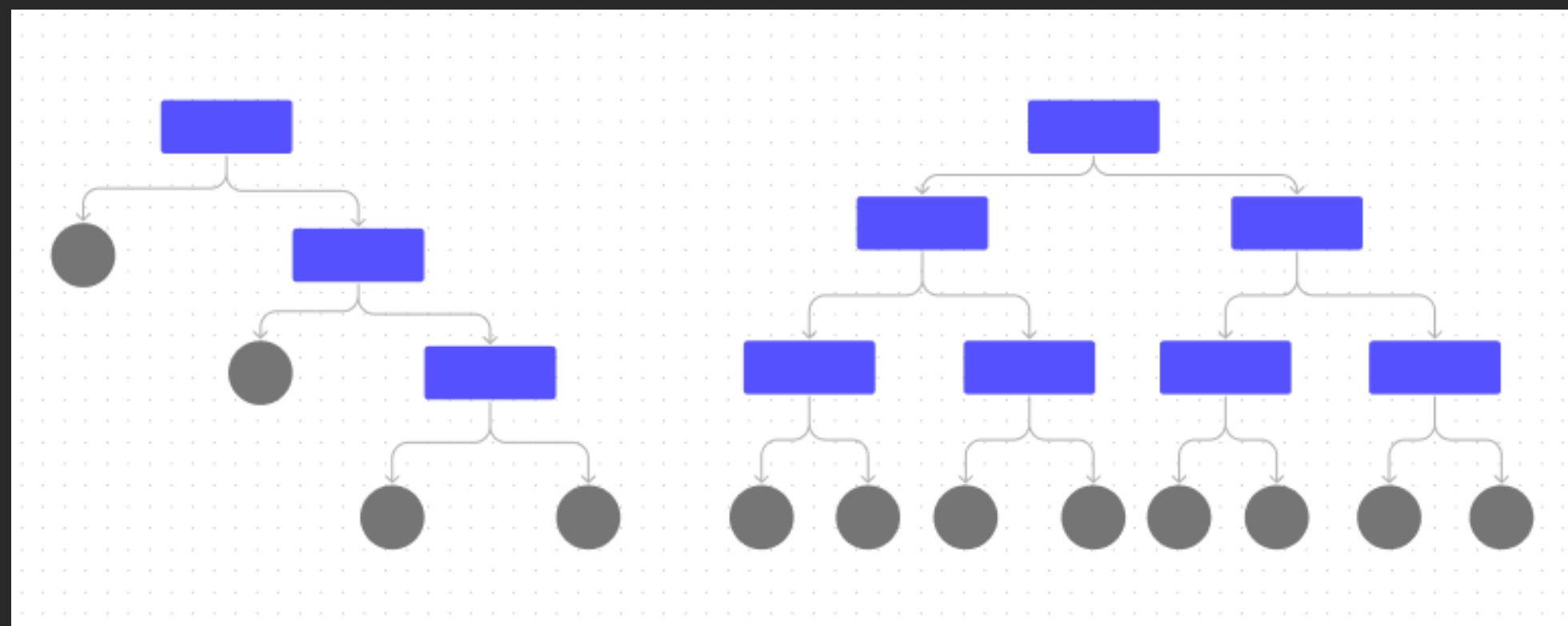


Base de Treino Final

- Muitos valores vazios
Principalmente em METAR e METAF> O que descarta modelos lineares
- Em nenhum momento o comportamento temporal das séries se mostrou ser um fator decisivo para predição de ELDT> O que descarta modelos como ARIMA, LSTM etc
- Variáveis categóricas importantes
Origem, Destino e Rota> Como rota tem uma cardinalidade alta, o ideal seria não haver nenhum tipo de transformação
- Rotas exibem comportamentos diferentes dependendo do período do dia> O modelo ideal deve ser capaz de assimilar esse comportamento complexo

CatBoost

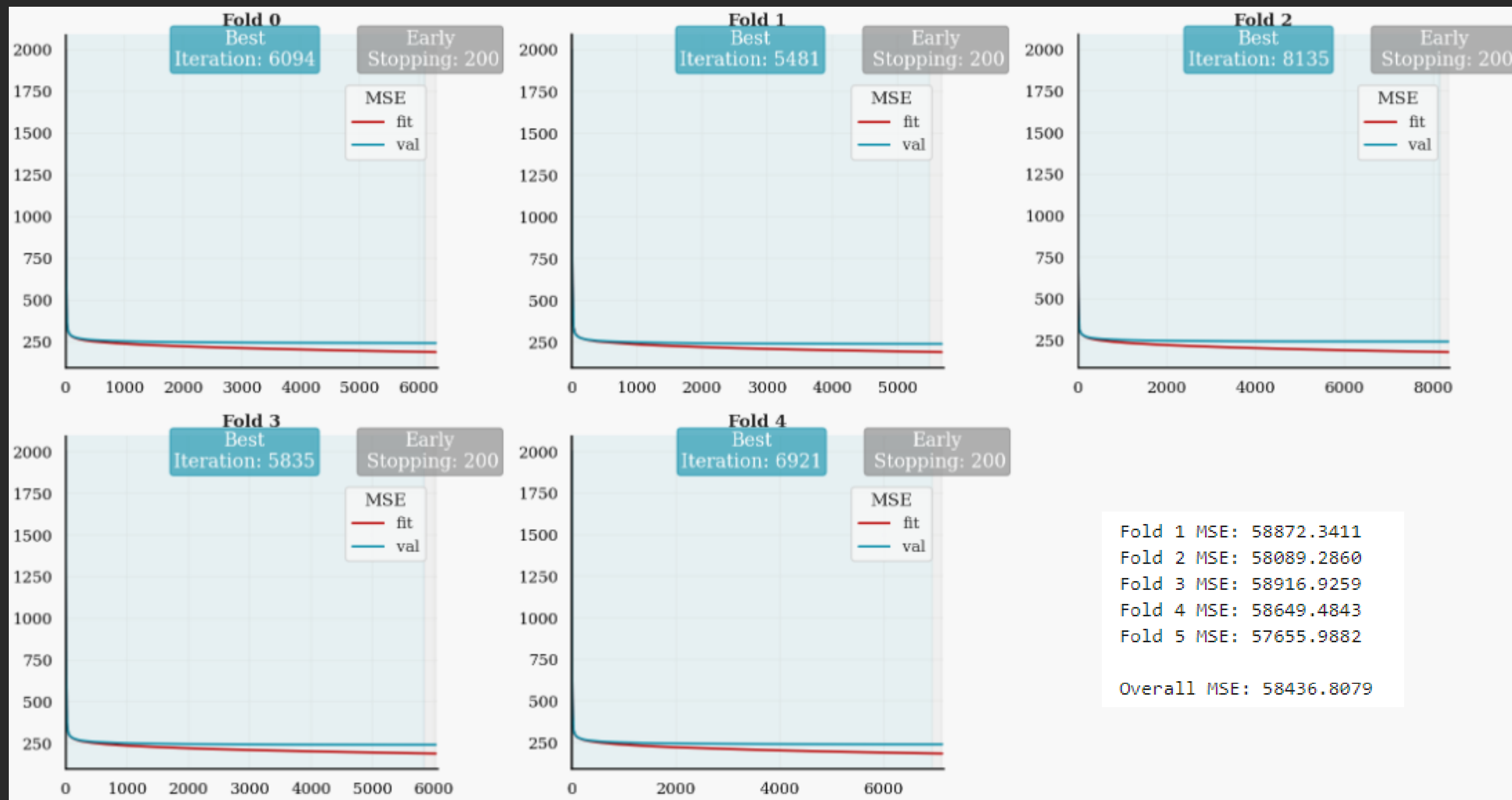
- Lida naturalmente com variáveis categóricas
- Oblivious Trees



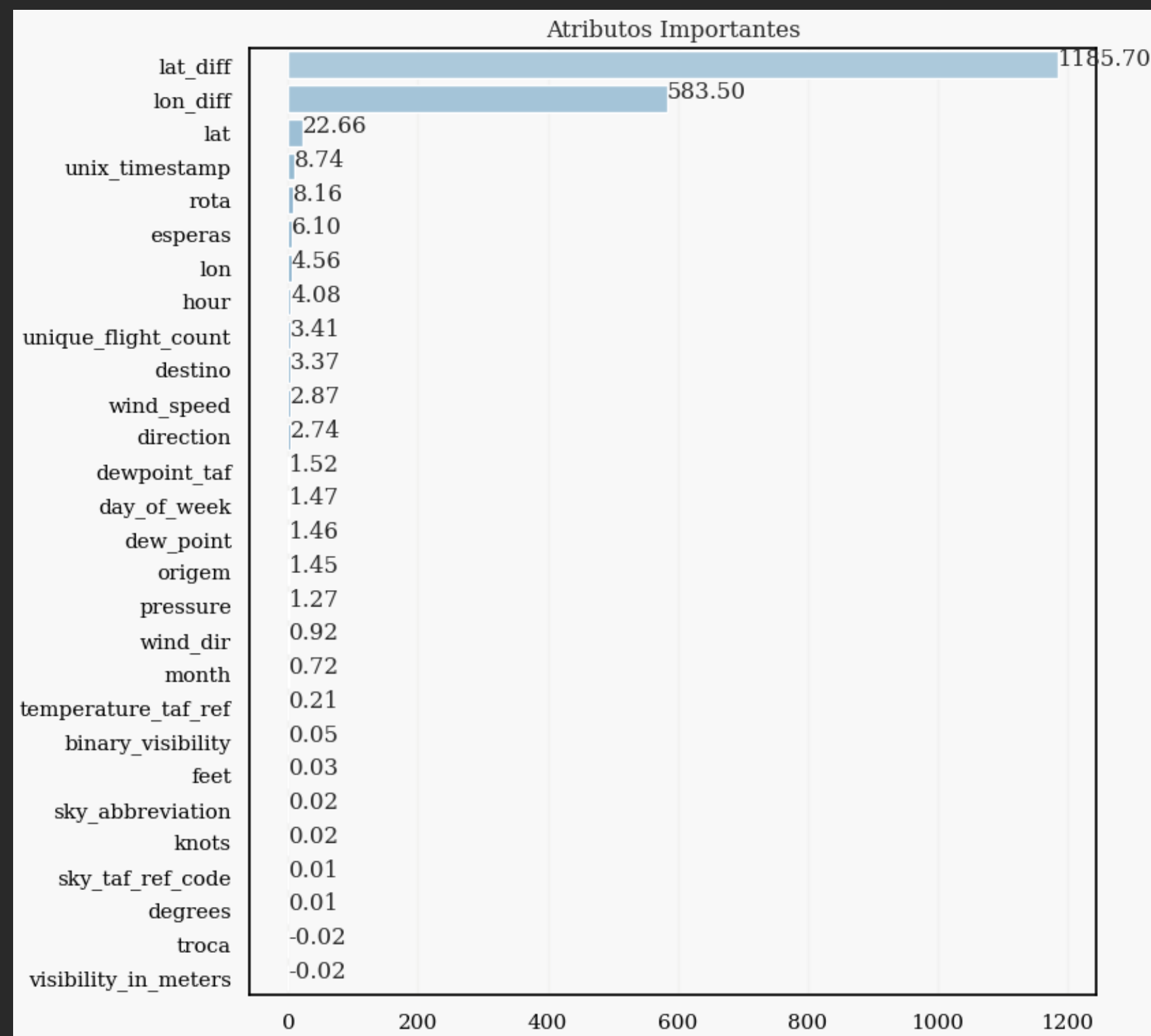
XGB, LGBM

CatBoost

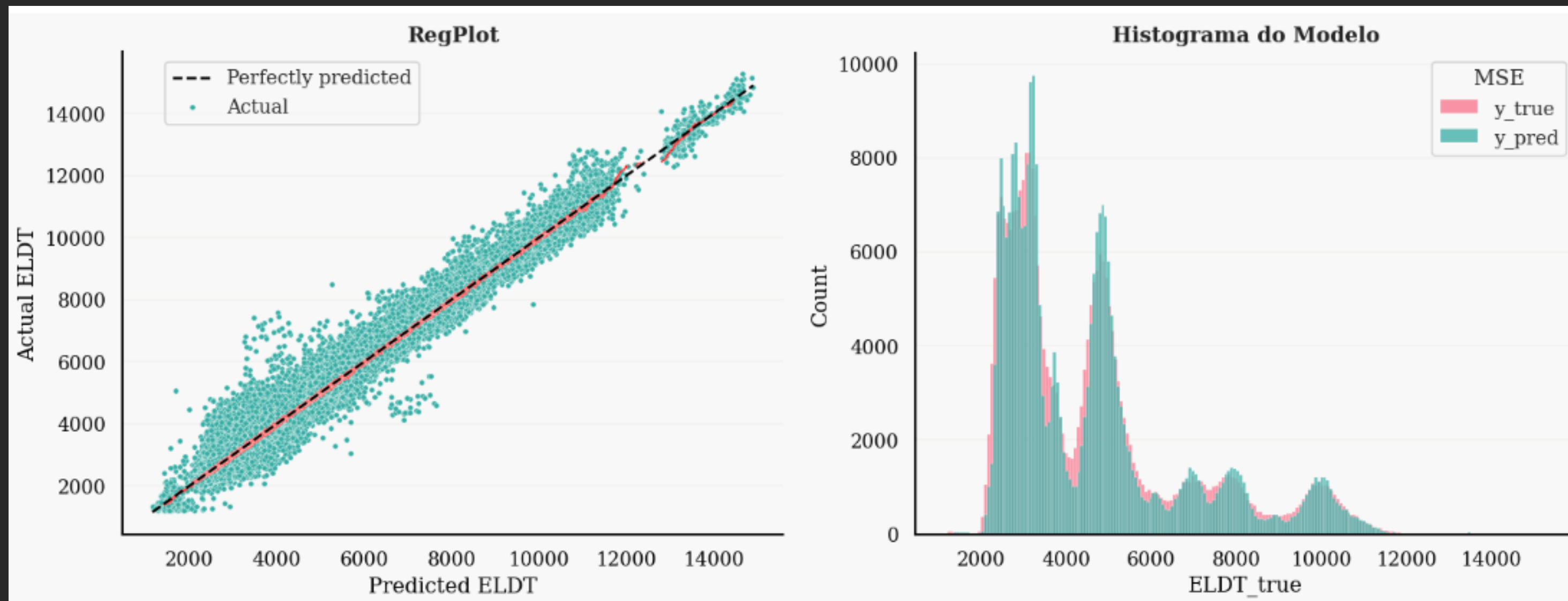
Treino



Atributos Importantes



Alguns Resultados



Alguns Resultados

	ELDT_predicted	ELDT_true	rota	month	abs_difference
141706	3469.814105	7407.0	SBCTSBPA	12	3937.185895
137331	3559.229203	7168.0	SBPASBCT	11	3608.770797
181108	4027.603232	7584.0	SBCTSBPA	8	3556.396768
10207	3627.156194	7175.0	SBCTSBPA	8	3547.843806
6757	3717.166699	7247.0	SBCTSBPA	7	3529.833301
193515	3285.157600	6810.0	SBCTSBPA	11	3524.842400
87154	3857.674632	7369.0	SBCTSBPA	12	3511.325368
144968	3251.335228	6655.0	SBPASBCT	12	3403.664772
207089	4038.840713	7440.0	SBCTSBPA	2	3401.159287
229173	3662.546500	7057.0	SBCTSBPA	7	3394.453500