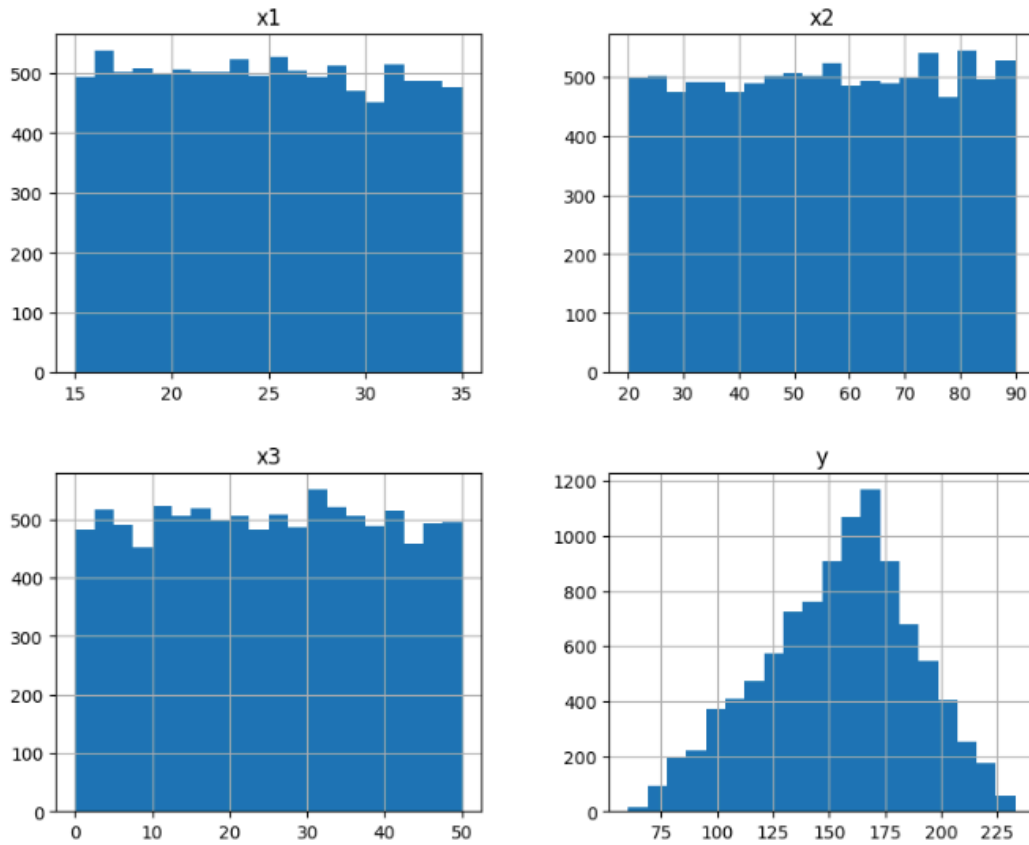


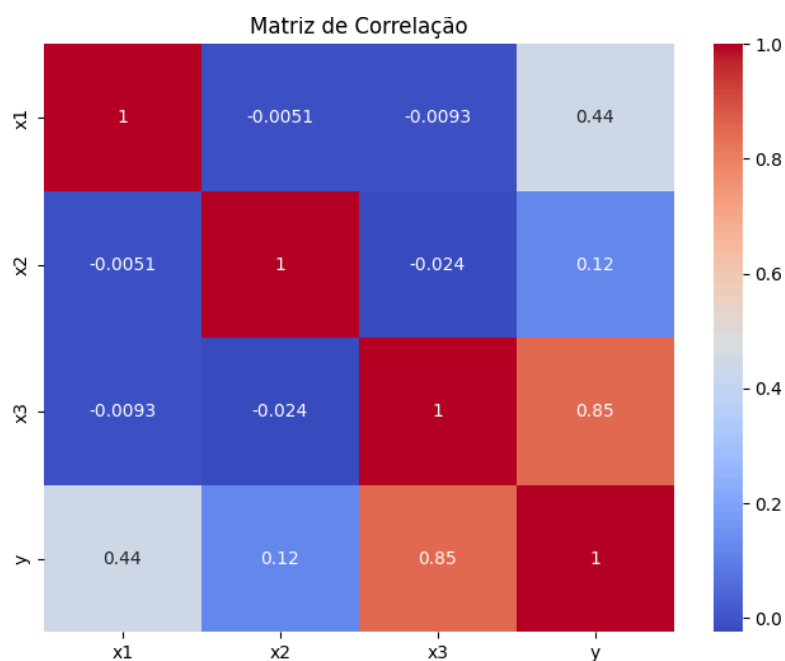
Alunos: Yann Azevedo Torres, Thaynara Damazio, Lucas Tucunduva, Thamirys Pinheiro, Raphael Loureiro e Luiz Felipe Medeiros.

1. Exploração de Dados

Histogramas das Variáveis



Acima conseguimos ver a distribuição dos valores de cada uma das variáveis do CSV.

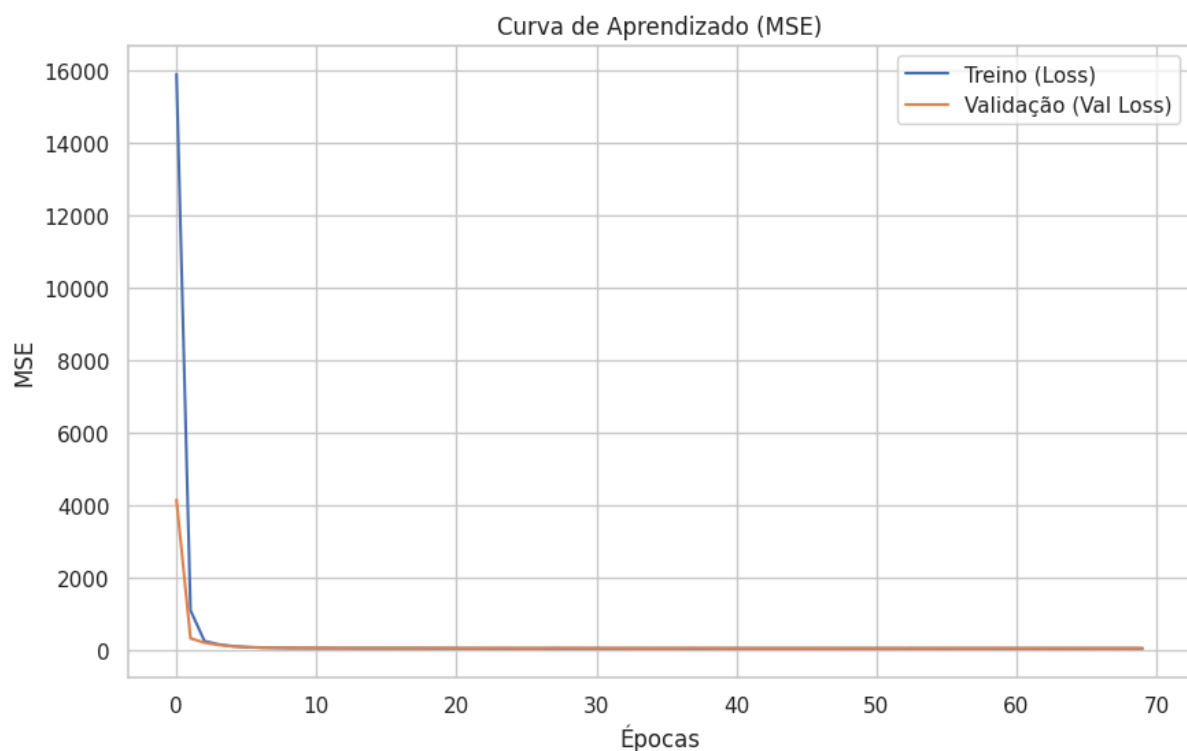


Sobre a Matriz de Correlação, a variável 'Ocupação' (x3) possui uma correlação positiva forte com o consumo de energia (y), sugerindo que quanto mais pessoas, maior o gasto. A temperatura (x1) também mostrou tendência de aumento no consumo.

2. Metodologia

Os dados foram divididos em 80% para treinamento e 20% para teste. Foi aplicada a padronização (StandardScaler) nas variáveis de entrada para garantir que a rede neural convirja mais rapidamente, já que as escalas de temperatura (graus) e ocupação (pessoas) são muito distintas. A rede utilizada foi um MLP com topologia 3-10-1, função de ativação ReLU na camada oculta e otimizador Adam.

3. Resultados e Treinamento do Modelo



O modelo convergiu antes das 300 épocas devido ao uso do Early Stopping, evitando o overfitting. Observa-se que a perda de validação acompanhou a perda de treino, indicando uma boa generalização.

4. Métricas de Avaliação

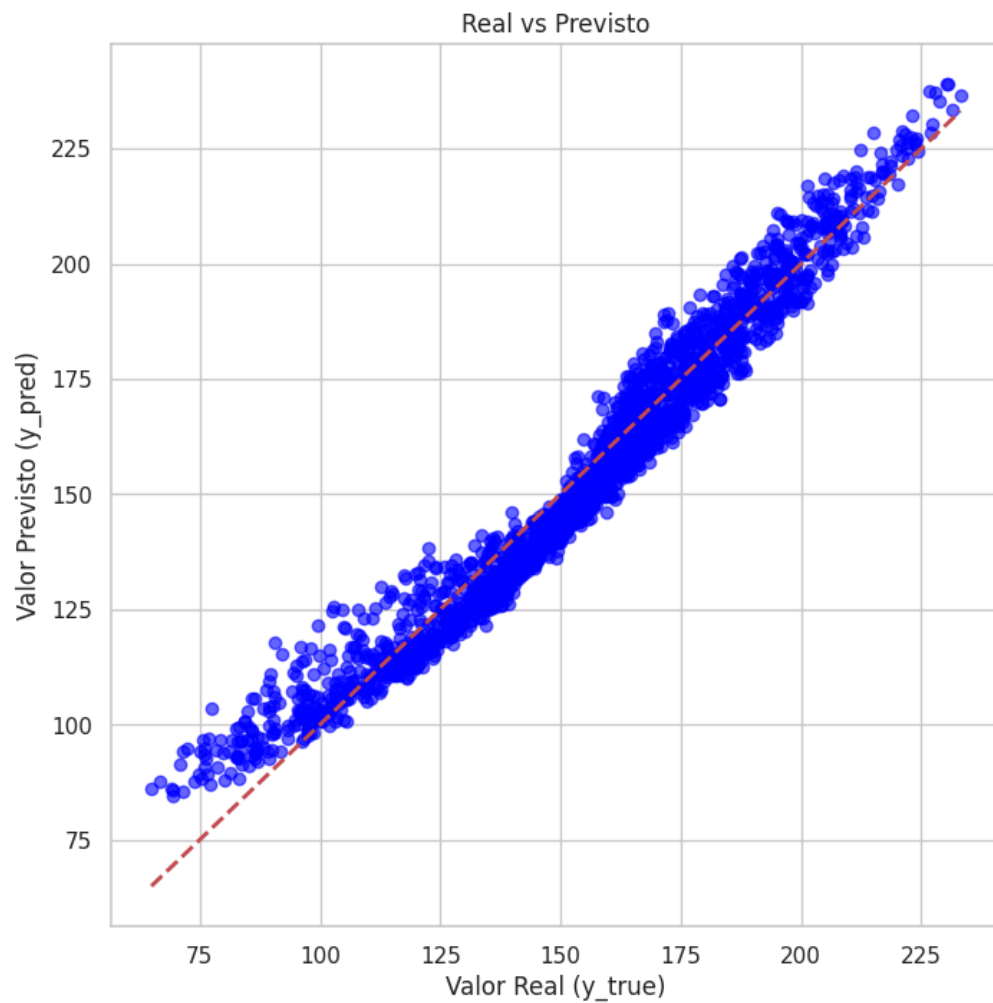
MSE (Erro Quadrático Médio): 9.8366

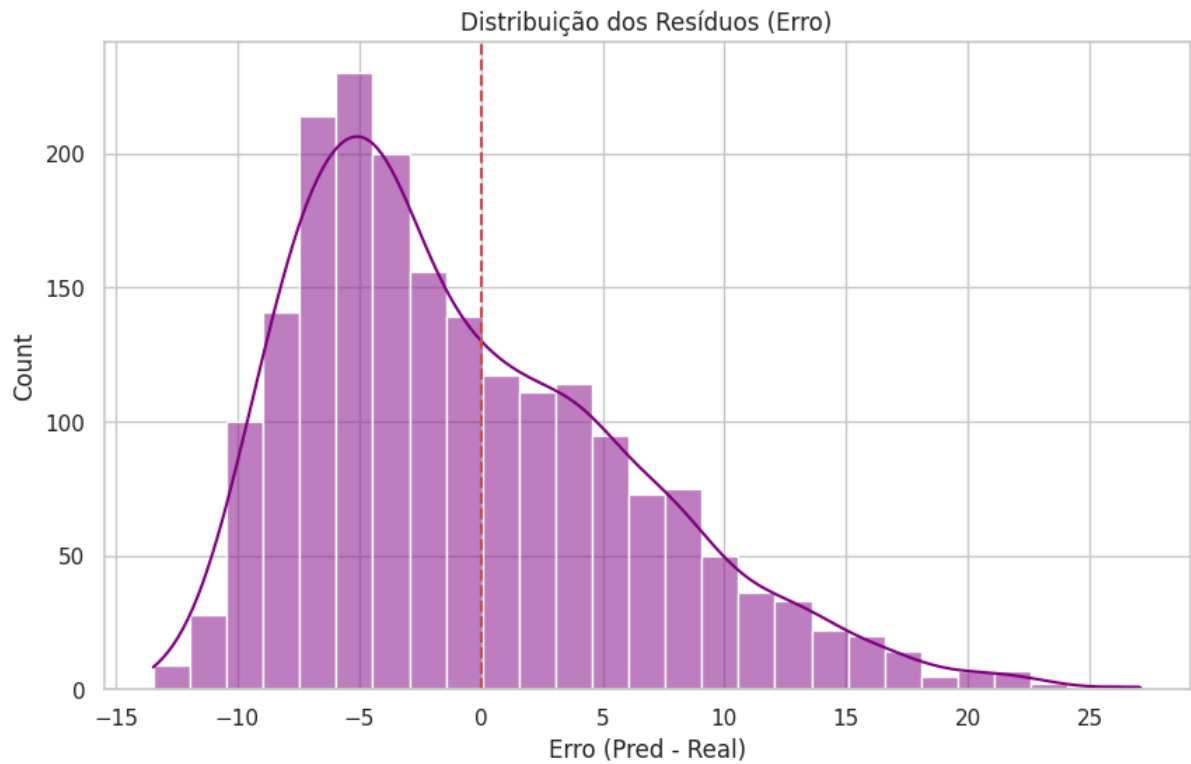
MAE (Erro Absoluto Médio): 2.4707

R^2 (Coeficiente de Determinação): 0.9907.

O modelo obteve um R^2 de 0.9907, que é bem próximo de 1.0 indicando que o modelo é excelente. O erro absoluto médio (MAE) foi de 2.4707 kWh, o que é considerado baixo para a escala de consumo do prédio.

5. Análise de Resíduos





A distribuição dos resíduos aproxima-se de uma normal centrada em zero, o que indica que o modelo não possui viés (bias) significativo e capturou bem a não-linearidade dos dados.