

**INSTITUTO  
FEDERAL**  
Ceará

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ -  
CAMPUS CANINDÉ  
ANÁLISE E DESENVOLVIMENTO DE SISTEMAS  
INTELIGÊNCIA ARTIFICIAL

# RELATÓRIO I

ISMAEL FERNANDES BRITO

CANINDÉ - CE  
2022

# 1 Introdução

Em muitos problemas práticos, há duas ou mais variáveis numéricas que parecem estar intrinsicamente relacionadas, sendo necessário analisar a natureza matemática dessa relação de maneira mais formal a fim de entender melhor o problema. A construção de modelos estatísticos é um problema de alta complexidade, existindo diversos fatores na sua concepção desde sua forma (linear ou não-linear) e quantidade de parâmetros do modelo.

Este trabalho busca avaliar o desempenho dos métodos (linear e não-linear) na aproximação da função da energia pela velocidade através do coeficiente de determinação  $R^2$  ajustado, o MSE (Mean Squared Error) e o RMSE (Root Mean Squared Error).

Todo o trabalho foi implementado utilizando as notas de aula e os códigos na linguagem R para rodar o conjunto de dados do aerogerador nos testes que este trabalho aborda.

## 2 Dados Utilizados

Os dados do problema tem como fonte o conjunto de observações da curva de potência de um aerogerador. O conjunto de dados tem 2250 observações, 1 entrada e 1 saída que são respectivamente a velocidade e a energia produzida pelo aerogerador.

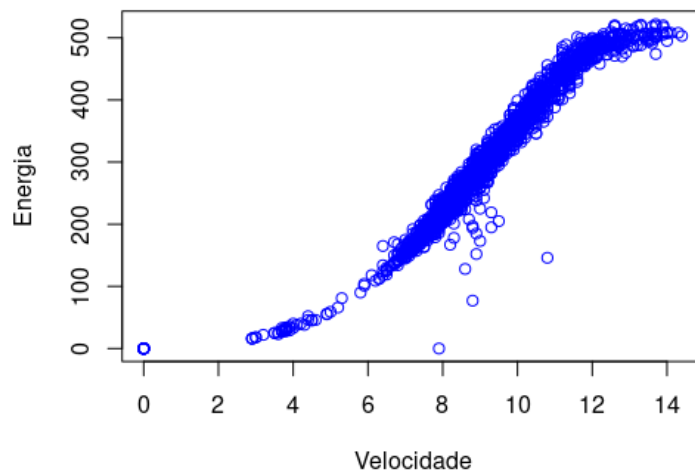


Figura 1: Gráfico de dispersão dos dados do aerogerador

## 3 Metodologia

Aplicando os três algoritmos ao conjunto de dados do aerogerador e utilizando a linguagem R para a implementação computacional para cada método de regressão com o seguinte procedimento abaixo:

- Carregamento dos dados;
- Divisão do conjunto de dados em treinamento (80%) e teste (20%);
- Cálculo do  $R^2$  e  $R^2$  ajustado.
- Cálculo do RMSE (Root Mean Squared Error).
- Cálculo do MAE (Mean Absolute Error).

## 4 Resultados

O valor do  $R^2$  ajustado para o modelo de regressão linear explica 93% da variância, ou seja, quando mais variância for explicada pelo modelo de regressão, mais próximos os pontos de dados estarão em relação à linha de regressão ajustada.

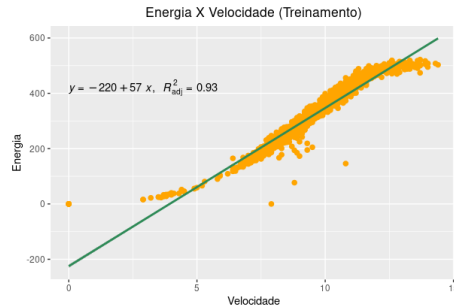


Figura 2: Gráfico de dispersão dos dados de treinamento do aerogerador e o gráfico da reta obtida pelo modelo de regressão linear

O valor do  $R^2$  ajustado para o modelo de regressão não-linear (polinomial) de grau 2 explica 94% da variância, ou seja, quando mais variância for explicada pelo modelo de regressão, mais próximos os pontos de dados estarão em relação à linha de regressão ajustada.

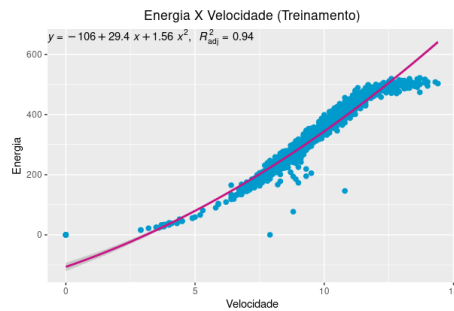


Figura 3: Gráfico de dispersão dos dados de treinamento do aerogerador e o gráfico da reta obtida pelo modelo de regressão não-linear (polinomial) de grau 2

O valor do  $R^2$  ajustado para o modelo de regressão não-linear (polinomial) de grau 3 explica 97% da variância, ou seja, quando mais variância for explicada pelo modelo de regressão, mais próximos os pontos de dados estarão em relação à linha de regressão ajustada.

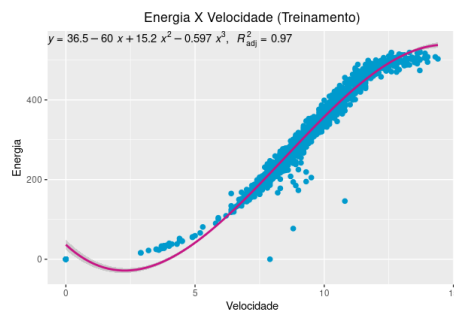


Figura 4: Gráfico de dispersão dos dados de treinamento do aerogerador e o gráfico da reta obtida pelo modelo de regressão não-linear (polinomial) de grau 3

Já em relação ao dados de teste do aerogerador no modelo de regressão linear obtivemos os seguintes resultados: RMSE (Root Mean Squared Error) de 32.0483577, o  $R^2$  de 0.9178335 e o MAE (Mean Absolute Error) de 20.0257428.

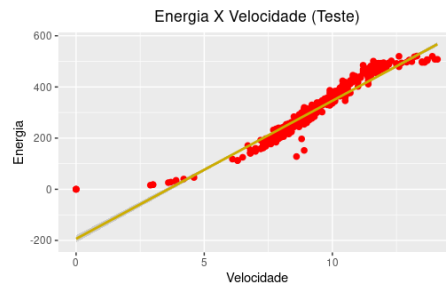


Figura 5: Gráfico de dispersão dos dados de teste do aerogerador com a reta obtida pelo modelo de regressão linear

Os dados de teste do aerogerador no modelo de regressão não-linear (Polinomial) de grau 2 obtivemos os seguintes resultados: RMSE (Root Mean Squared) de 26.5115286, o  $R^2$  de 0.9420451 e o MAE (Mean Absolute) de 18.4607481.

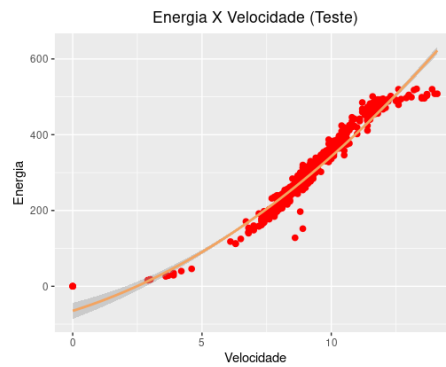


Figura 6: Gráfico de dispersão dos dados de teste do aerogerador e o gráfico da reta obtida pelo modelo de regressão não-linear (polinomial) de grau 2

Os dados de teste do aerogerador no modelo de regressão não-linear (Polinomial) de grau 3 obtivemos os seguintes resultados: RMSE (Root Mean Squared) de 18.5174372, o  $R^2$  de 0.9715025 e o MAE (Mean Absolute) de 12.8059753.

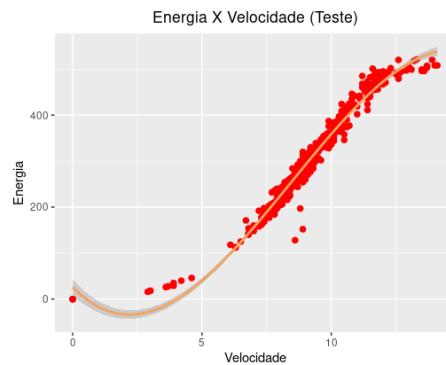


Figura 7: Gráfico de dispersão dos dados de teste do aerogerador e o gráfico da reta obtida pelo modelo de regressão não-linear (polinomial) de grau 3

## 5 Conclusão

Este estudo teve como meta a aplicação de modelos de regressão com o objetivo de obter a melhor curva que se ajusta aos dados do aerogerador, sendo utilizados os modelos linear e não-linear (Polinomial) de grau 2 e 3.

Pelo gráfico de dispersão da figura 1, vemos que os dados do aerogerador tem um comportamento não-linear, então não vale a pena avaliar o modelo de regressão linear. Quando partimos para o modelo de regressão não-linear (Polinomial), também não vale a pena avaliar um modelo polinomial de grau 2, porque percebemos duas concavidades ou curvaturas no gráfico, uma no início e outra no fim. Por tanto, vamos partir nossa análise com o modelo polinomial de grau 3, obtemos um coeficiente de determinação  $R^2$  ajustado alto, porém na faixa de 0 a 5, o modelo revê potências negativas, o que não faz sentido, ou seja, o modelo está dizendo que o aerogerador consome energia ao invés de gerá-la.

Conclui-se os três modelos utilizados neste estudo não apresentou resultados satisfatório de encontrar a melhor curva que representasse os dados do aerogerador.