

UMA COMPARAÇÃO DE 3 TÉCNICAS DE PREVISÃO PARA 5 PASSOS À FRENTE

Carlos F. Oliveira

Escola Politécnica de Pernambuco – Universidade de Pernambuco (UPE)
– Recife – PE – Brasil

carlos.fe.oliveira@gmail.com

Abstract. *This paper compares the use of 3 different forecasting techniques used in 2 databases. Predictions are made to predict 5 steps forward, using direct and iterative forecasting methods to compare the results and see the positives and negatives of each method.*

Resumo. *Este artigo compara o uso de 3 técnicas de previsão diferentes, utilizados em 2 bases de dados. As previsões são feitas para prever 5 passos à frente, e para isso são utilizados métodos de previsão direta e iterativa, afim de se comparar resultados e ver os pontos positivos e negativos de cada método.*

1. Introdução

Os estudos em séries temporais tem crescido rapidamente nos últimos anos, com a evolução da computação e da programação a previsão deixou de ser utilizada apenas por empresas grandes e universidades para ser utilizado pelo usuário comum. Diversos métodos e abordagens foram criados para realizar a mesma previsão, entretanto cada método possui seus pontos positivos e negativos e podem ser bons para previsão de um certo tipo de dado, mas ruins para outro. Logo, no campo da previsão das séries temporais não existe verdade absoluta sobre qual a melhor abordagem, e nesse artigo busca-se a comparação em cima de bases de dados distintas e das técnicas de previsão utilizando Rede Neural Artificial (RNA), a máquina de suportes de vetores de regressão, do inglês *Support Vector Regression (SVR)* e a Regressão Linear.

2. Revisão de Literatura

2.1. Regressão Linear

2.1.1. Regressão Linear Simples

A regressão linear simples é um modelo que possui uma única variável independente x que tem uma relação com a variável dependente y . A equação para obtê-la é dada por:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

Onde β_0 e $\beta_1 x$ são os coeficientes de regressão estimados e ε é um erro que conta a discrepância entre o valor predito na Equação (1) e o valor observado. O valor predito na Equação (1) é dado por:

$$\bar{y} = \bar{\beta}_0 + \bar{\beta}_1 x \quad (2)$$

Onde \bar{y} é o valor predito, e $\bar{\beta}_0$ e $\bar{\beta}_1$ são os coeficientes de regressão estimados.

2.1.2. Regressão Linear Múltipla

A regressão linear múltipla é a generalização da regressão linear simples, permitindo

mais de uma variável independente. Segue abaixo a equação para obtê-la.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots + \beta_p x_p + \varepsilon \quad (3)$$

Onde y é a variável de resposta, $x_1, x_2, x_3, \dots, x_p$ são as variáveis independentes sendo p o número de variáveis, $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \dots + \beta_p$ são os coeficientes de regressão e ε é um erro que conta a discrepância entre o valor predito na Equação (1) e o valor observado. O valor predito na Equação (3) é:

$$\bar{y} = \bar{\beta}_0 + \bar{\beta}_1 x_1 + \bar{\beta}_2 x_2 + \bar{\beta}_3 x_3 \dots + \bar{\beta}_p x_p \quad (4)$$

Onde \bar{y} é o valor predito, e $\bar{\beta}_0, \bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3 \dots \bar{\beta}_p$ são os coeficientes de regressão estimados.

2.2. Redes Neurais Perceptron Multicamadas

As redes neurais artificiais foram desenvolvidas com o objetivo de criar um modelo científico computacional de aprendizado similar ao cérebro humano. Elas são capazes de aprender os principais padrões de informação dentro de um domínio de informação multidimensional. Para o seu funcionamento não é necessário haver uma informação detalhada sobre o sistema, pois ela funciona como um modelo caixa preta. Ela aprende a relação entre os parâmetros de entrada e as variáveis controladas e não controladas através do estudo dos dados previamente registrados. De acordo com, as redes neurais possuem a vantagem de poder executar tarefas que um programa linear não pode e a desvantagem de ter um tempo de processamento elevado quando a rede neural é grande.

As redes neurais perceptron multicamadas (MLP) são constituídas de neurônios que interagem entre si utilizando conexões ponderadas. Elas obtiveram bons resultados em diversos problemas dentre estes, alguns considerados difíceis. Basicamente uma rede MLP é composta por uma camada de entrada, que consiste em um conjunto de unidades sensoriais (nós de fonte), uma ou mais camada(s) escondida(s) e uma camada de saída, sendo que as duas últimas camadas possuem capacidade computacional. O sinal de entrada propaga-se de camada em camada sempre para frente. A camada de entrada é onde reside os dados de entrada e é composta por unidades sensoriais. Sua função é encaminhar os dados à primeira camada oculta. A(s) camada(s) oculta(s) tem a função de receber os dados da camada de entrada, processá-los e encaminhá-los à camada de saída. A camada de saída tem a função de fornecer a resposta propagada pela camada oculta.

2.3. SVR

O algoritmo máquina de vetores de suporte para regressão (SVR) é um algoritmo de aprendizado de máquina supervisionado que tem como objetivo encontrar um modelo de decisão para representar a relação entre os atributos e o seu objetivo. Ele se baseia no princípio indutivo da minimização do risco estrutural, que procura minimizar um limite superior da generalização do erro que consiste na soma do erro de treinamento e um nível de confiança. Esse princípio faz com que muitos problemas no aprendizado de máquina tenham melhores desempenhos de generalização com o SVR do que com outras técnicas. O treinamento do SVR é equivalente a resolver um problema restrito de programação quadrática de forma linear de modo que a solução do algoritmo é sempre globalmente ótima e única, o que evita o risco de mínimos locais. A solução para o problema no SVR é dependente apenas de um subconjunto de pontos de dados de treinamento, chamados de vetores de suporte. A complexidade de tempo de treinamento em qualquer parte do SVR é quadrática ou cúbica em relação ao número de amostras de treinamento, uma das poucas desvantagens desse algoritmo. De acordo com a literatura, o SVR tem se mostrado um algoritmo muito eficiente para resolver problemas gerais de regressão.

Basicamente o objetivo do SVR consiste em encontrar uma função que aproxime da melhor maneira possível os pontos de treinamento com a meta de minimizar o erro de previsão. Dois parâmetros são especificados pelo usuário o ε e o C e a calibração deles é peça fundamental para a qualidade da previsão. O ε define um tubo em torno da função de regressão que controla o grau de aproximação da função aos dados de treinamento. E C é a constante de regularização, ela determina o limite superior de *tradeoff* entre o erro de previsão e o nivelamento do tubo ε .

2.4. Erro Quadrático Médio

O erro quadrático médio (EQM) também pode ser usado como uma medida do erro de previsão. O EQM é determinado somando os erros de previsão ao quadrado e dividindo pelo número de erros usados no cálculo. O erro quadrático médio pode ser expresso pela seguinte equação:

$$EQM = \frac{\sum_{t=1}^n e_t^2}{n}$$

3. Metodologia

Para esse estudo foram utilizadas 2 bases de dados cedidas pelo professor da disciplina, uma base mostra o número de passageiros por avião da Airlines, e a outra mostra o consumo de energia elétrica. Nas técnicas de RNA e SVR foram utilizados dois métodos diferentes para a previsão de 5 passos a frente, o método iterativo e o direto. Já o método de Regressão Linear foi utilizado apenas o método direto de previsão.

O método direto consiste em treinar, desde o começo, a previsão para já prever 5 passos a frente do seu tempo, porém esse método deixa uma lacuna das previsões para as 4 previsões anteriores. Para suprir essa necessidade é necessário fazer mais 4 modelos de previsão de 1, 2, 3 e 4 passos. Isso demanda um poder computacional maior para fazer a previsão que acaba tendo que treinar e fazer 5 modelos de previsões distintas. O lado positivo desse método é que tende a ter um erro menor, já que modelo já é treinado e a previsão é feita de forma direta.

O método iterativo, diferentemente do método direto, só treina seu modelo para prever 1 passo a sua frente, ele consiste, neste caso, em fazer 5 previsões seguidas de 1 passo, utilizando o resultado da sua previsão anterior como uma entrada para a previsão do próximo passo. Alcançando, assim, a previsão desejada de 5 passos a frente. Esse modo necessita de menor poder computacional, já que só precisa treinar um único modelo. Porém, por ele utilizar a sua própria previsão como uma entrada para a próxima previsão, se uma previsão é feita com um erro, a próxima vai acumular o erro da anterior, fazendo com que sua previsão fique menos precisa.

Para realizar as previsões foi utilizado a linguagem Python, o código e a compilação foi realizada no ambiente do Spyder. O programa calculava o erro quadrático médio e plotava o gráfico do valor real e da previsão para comparação da qualidade da previsão, caso o resultado não fosse satisfatório, os “coeficientes” de cada técnica era alterado até encontrar um resultado satisfatório.

Para a previsão utilizando Rede Neural Artificial, foi utilizado uma biblioteca do Python chamada MLP Regressor, que gerava o modelo e treinava com base nos dados utilizados. Na máquina de suporte de vetores de regressão, foi utilizado uma busca cega para achar os melhores parâmetros para a previsão.

4. Resultados

O estudo consiste na comparação de diversas técnicas de previsão utilizando mais de um método, então os resultados das previsões (azul) foram plotados juntos com seus

resultados reais (vermelho) e os próximos 5 resultados (laranja), o erro quadrático médio também foi calculado para depois comparar todos juntos.

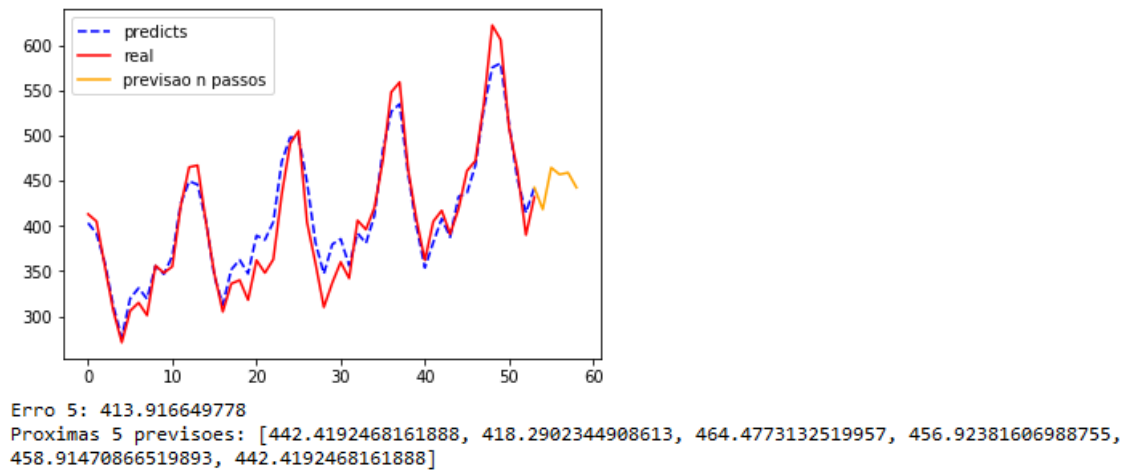


Figura 1: Rede Neural de previsão direta - Airlines

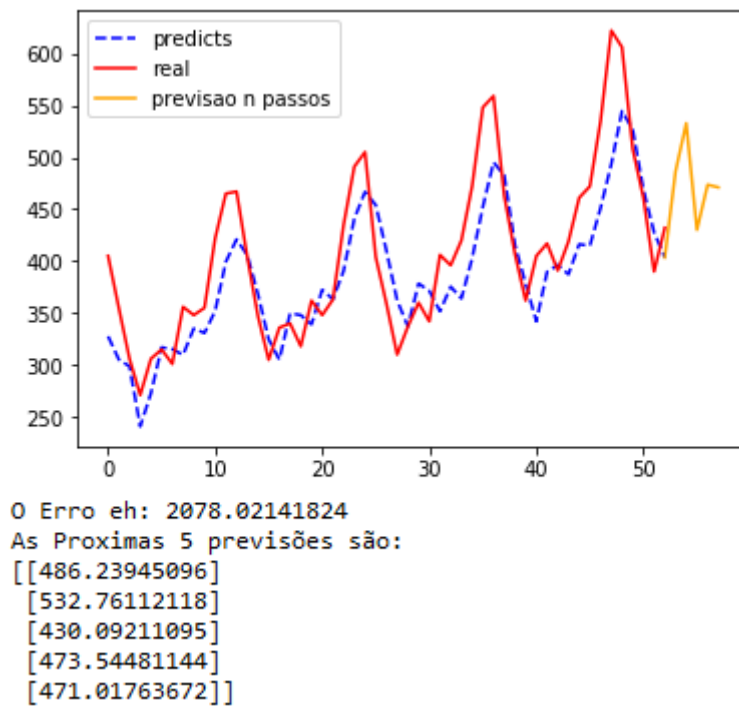
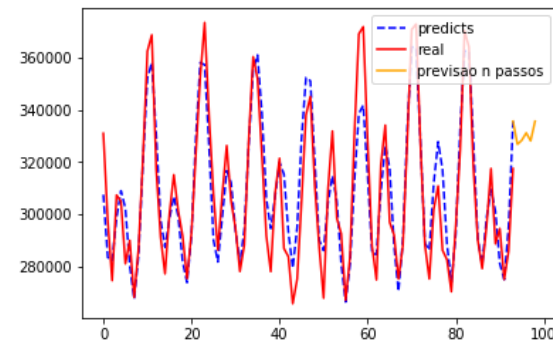
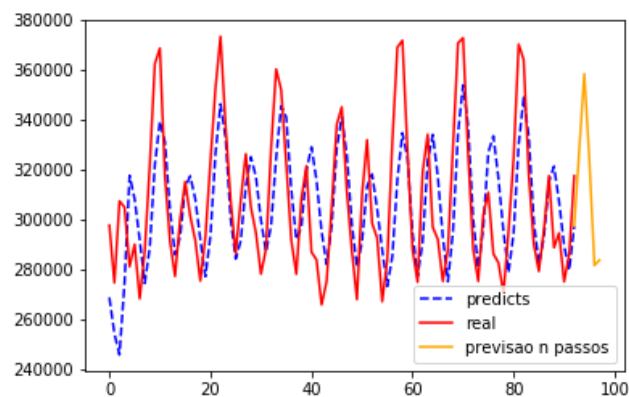


Figura 2: Rede Neural de previsão iterativa - Airlines



Erro 5: 139574849.433
 Proximas 5 previsoes: [335586.49725578923, 326739.7739992088, 328321.1542671136, 331131.26738910435, 328058.63175667205, 335586.49725578923]

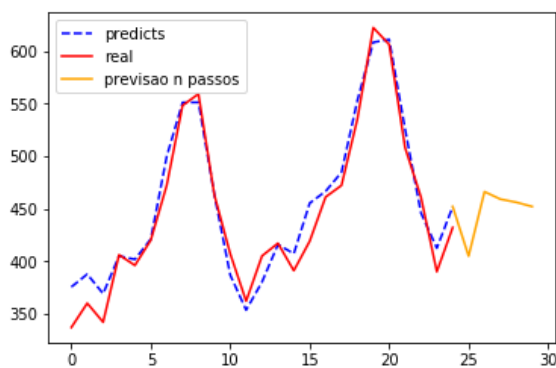
Figura 3: Rede Neural de previsão direta - Eletric



O Erro eh: 532390571.182
 As Proximas 5 previsões são:
 [[327900.21347769]
 [358371.86264331]
 [323924.60192244]
 [281581.3793472]
 [283630.72761091]]

Figura 4: Rede Neural de previsão iterativa - Eletric

Para a RNA direta foi utilizado um número de 12 camadas escondidas e já obteve um resultado satisfatório, já na iterativa foi utilizada 100 camadas escondidas para melhorar o erro, mas ainda não chegou a um resultado como na direta.



Proximas 5 previsoes: [451.8836696842531, 404.78418000653085, 466.02891715260694, 458.9407498940596, 455.9902055754038, 451.8836696842531]
 Erro 5: 346.511923867

Figura 5: SVR de previsão direta - Airlines

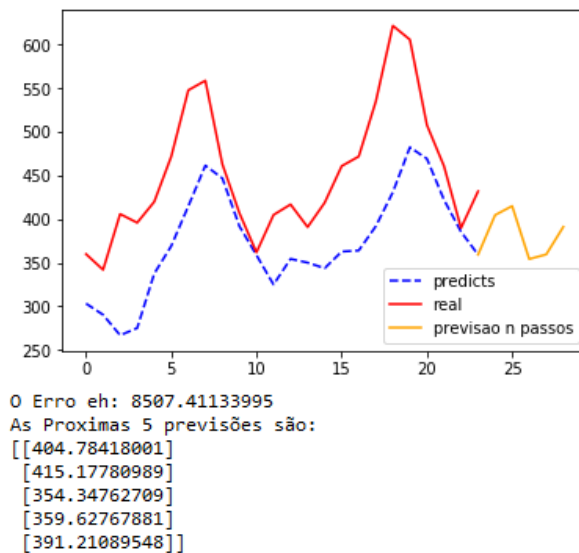


Figura 6: SVR de previsão iterativa - Airlines

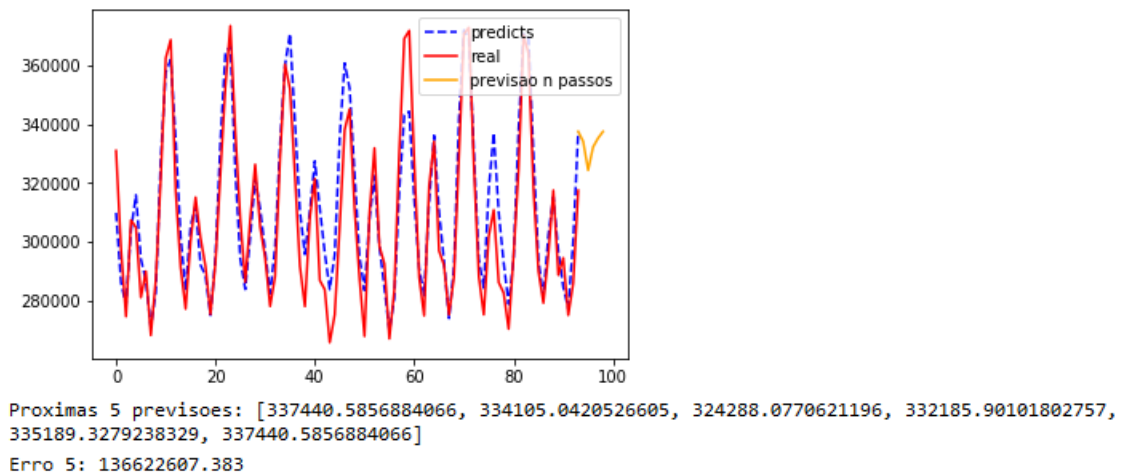


Figura 7: SVR de previsão direta - Eletic

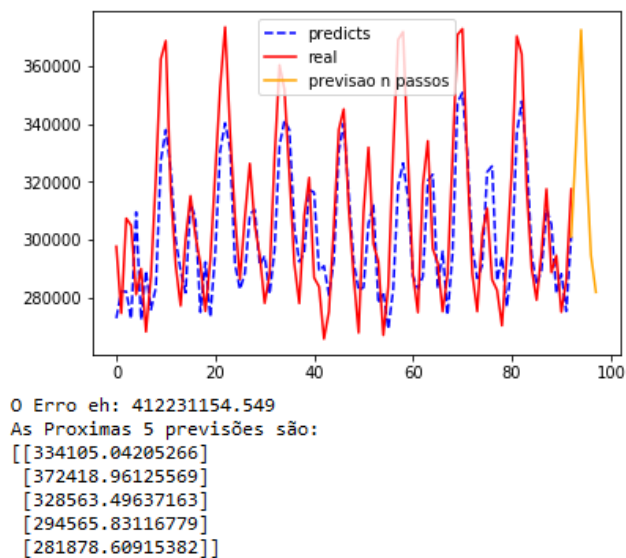
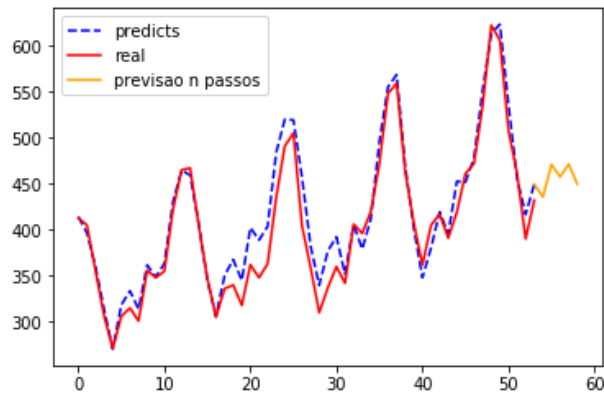


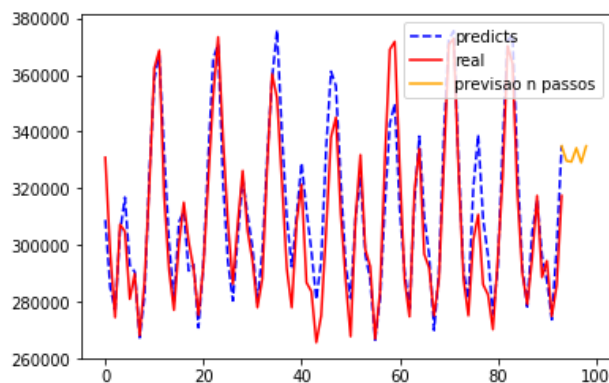
Figura 8: SVR de previsão iterativa - Eletic



Erro 5: 439.916507421

Proximas 5 previsoes: [449.516912435583, 435.5817561043353, 470.85108317329883, 457.4702151052921, 471.32604219906756, 449.516912435583]

Figura 9: Regressão Linear de previsão direta – Airlines



Erro 5: 130655124.923

Proximas 5 previsoes: [334943.78386752855, 329710.02216433256, 329394.28049428924, 334231.1307727983, 329147.3172167586, 334943.78386752855]

Figura 10: Regressão Linear de previsão direta - Eletric

	Rede Neural		SVR		Regressão Linear
	Direta	Iterativa	Direta	Iterativa	Direta
Airlines	413,917	2.078,021	346,512	8.507,411	439,916
Eletric	139.574.849,433	532.390.571,182	337.440,586	412.231.154,549	130.655.124,923

Tabela 1: Erro de cada previsão

5. Conclusão

Os resultados na maioria dos casos foram satisfatórios em todos as técnicas de previsão, foi observado que o método iterativo tende a ter um erro maior e ser mais rápido, que o método direto. A RNA é muito dependente da sua semente de ativação, embora tenha resultados bons na maioria das vezes, em alguns momentos aconteceu com um erro muito grande.

O método de Regressão Linear, apesar de ser um método linear, apresentou nesses

dados resultados muito bons comparado com o SVR e RNA.

Apesar de o resultado já ser bom, todos os métodos ainda podem ser melhorados utilizando métodos de otimização, o uso de métodos como enxame de partículas, no inglês *Particle Swarm Optimization (PSO)*. Teria resultados significativos nas previsões.

Referências

Montgomery, D. C., Peck, E. A., e Vining, G. G. (2015). *Introduction to linear regression analysis*. John Wiley & Sons.

Vasconcelos, Leonardo; João, Meirelles (2017). Uma comparação de técnicas de regressão para previsão de consumo de energia residencial no cenário nacional. Acessado em: 13/12/2018 em: <http://www.sbp2017.iltc.br/pdf/169301.pdf>

Braga, A.P.; Carvalho, A.P.L; Ludermir, T.B; Redes Neurais Artificiais: Teoria e Aplicações, LTC 2ª edição.

Coppin, B. Inteligência Artificial, LTC

Russel, S.; Norvig, P. Inteligência Artificial