



**CT-221**

**Laboratório 1BIM**

**Predição na Bolsa de Valores com metodologia  
NARMAX e arquitetura MLP**

Alunos

**Giuseppe Vicente Batista**

**Gustavo Gomes Santiago**

**Gustavo Pádua Beato**

**Pedro Pinheiro Borges**

**Victoria Maria de Araujo Ferreira Felix**

Professor

**Paulo Marcelo Tasinaffo**

1 de outubro de 2023

**Instituto Tecnológico de Aeronáutica (ITA)**

**Departamento de Ciência da Computação**

# 1 Introdução

Um Perceptron Multicamadas (MLP) é um tipo de rede neural com camadas interconectadas, incluindo entrada, ocultas e saída. Treinado por algoritmos como retropropagação, ajusta os pesos das conexões para aprender padrões complexos em dados, sendo amplamente usado em tarefas de classificação e regressão. A Figura 1 ilustra a arquitetura de uma rede neural deste tipo com uma camada interna não-linear e uma camada de saída linear. Pode ser demonstrado matematicamente que tal arquitetura de rede é um aproximador universal de funções.

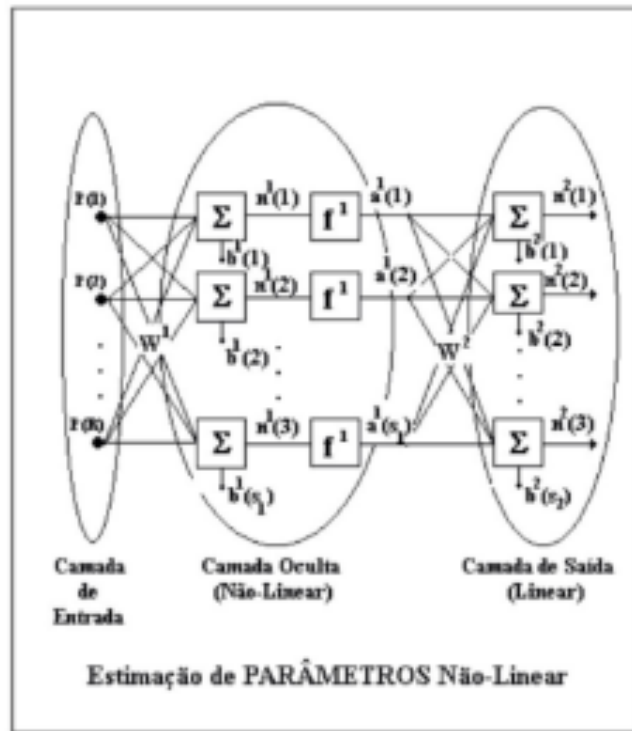


Figura 1: Arquitetura detalhada de uma rede Multilayer Perceptron com uma camada interna não-linear e uma camada de saída puramente linear.

A Metodologia NARMAX (Nonlinear AutoRegressive Moving Average with eXogenous inputs) é uma abordagem para modelar sistemas dinâmicos não lineares. Utiliza séries temporais para expressar saídas em termos de combinações não lineares de entradas e saídas passadas, usando técnicas de otimização para identificar estruturas e parâmetros. Aplicável em controle de processos e previsão de séries temporais complexas. A Figura 2 exibe graficamente como é utilizada a metodologia NARMAX para representar ou estimar a dinâmica de uma planta do mundo real. A planta pode ser qualquer sistema que tenha comportamento parecido com sistemas dinâmicos que podem ser regidos por equações diferenciais ordinárias. Exemplos de uma planta seriam: (i) funcionamento da bolsa de valores, (ii) movimentos de atitude de satélites, que devem ser controlados de forma autônoma em órbitas ao redor da terra e (iii) a dinâmica da variação de vazões de rios no decorrer das estações do ano. Observando a Figura 2 percebe-se que através de um aprendizado supervisionado a rede neural é treinada para

aprender o comportamento real da planta dentro de um erro aceitável. Teoricamente, isso será sempre possível através de qualquer tipo de rede com arquitetura feedforward (e.g., MLP, RBF, SVM, rede de Wavelets, sistemas Neuro-fuzzy, etc.).

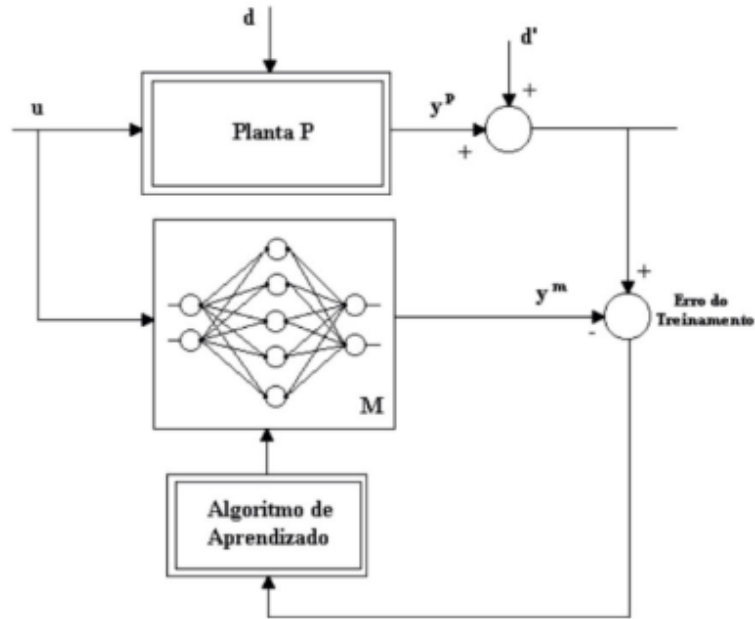


Figura 2: Metodologia NARMAX utilizada para representação de sistemas dinâmicos através de uma rede neural Multilayer Perceptron.

## 2 Objetivo

O presente laboratório teve como objetivo resolver o problema de predição de ações na Bolsa de Valores utilizando a metodologia NARMAX projetada com uma arquitetura de rede neural do tipo Multilayer Perceptron. Foram utilizados dados históricos de fechamento diário de ações de 3 empresas, por um período de aproximadamente 4 anos, a saber: Petrobras (PETR3.SA), Embraer (EMBR3.SA) e Vale do Rio Doce (VALE3.SA). A meta foi prever os últimos 3 meses das ações da Petrobras com diferentes abordagens de treinamento e configuração de camadas internas, comparando-as em termos de seus resultados.

## 3 Metodologia

Foram montadas três redes MLP na abordagem vetorial e considerando também variáveis exógenas. Foram utilizadas uma variável principal e duas variáveis exógenas, em cada uma destas três redes MLP. A Figura 3 ilustra apenas uma destas três redes neurais, que é a que foi utilizada para prever as ações da Petrobras. Nesta figura, tem-se as dez primeiras entradas na camada de entrada da rede (entrada vetorial principal da variável  $y_1$ ) e mais outras 20 entradas das variáveis exógenas ( $y_2$  e  $y_3$ ). Sendo assim, foram utilizadas três ações (PETR3, EMBR3 e VALE3) como entradas da primeira rede neural ( $y_1$ ,  $y_2$  e  $y_3$ ) e apenas um histórico de ações na saída da rede ( $y_1$ ). Na entrada desta primeira rede MLP  $y_1(t)$  é o valor presente,  $y_1(t-9.DT)$ ,

$y1(t-8.DT), \dots, y1(t-DT)$  são os valores atrasados ou do passado. Na saída desta primeira rede  $y1(t+DT), y1(t+2.DT), \dots, y1(t+10.DT)$  são os valores adiantados ou do futuro. Para correto funcionamento do modelo, foram utilizadas duas outras redes, a saber: (i) uma com variável principal  $y2$  e variáveis exógenas  $y1$  e  $y3$ , que foi utilizada para prever ações da Embraer e (ii) outra rede com variável principal  $y3$  e variáveis exógenas  $y1$  e  $y2$ , que foi utilizada para prever ações da Vale do Rio Doce. A função  $y1(t)$  se refere somente a uma série temporal, neste caso, as variações temporais das ações da Petrobras num período de quatro anos com coleta diária ( $DT=1$  dia). Assim, fez-se uma previsão diária ( $DT= 1$  dia) para um horizonte de 90 dias (3 meses). Desta forma, os primeiros três anos e nove meses desta série temporal foram utilizados no treinamento da rede neural. Assim, fez-se a previsão para os últimos três meses, desta série temporal, utilizando as três redes já treinadas previamente. As redes neurais foram treinadas com diversas abordagens, todas com modelo NARMAX com variáveis exógenas, conforme detalhado na Tabela 1.

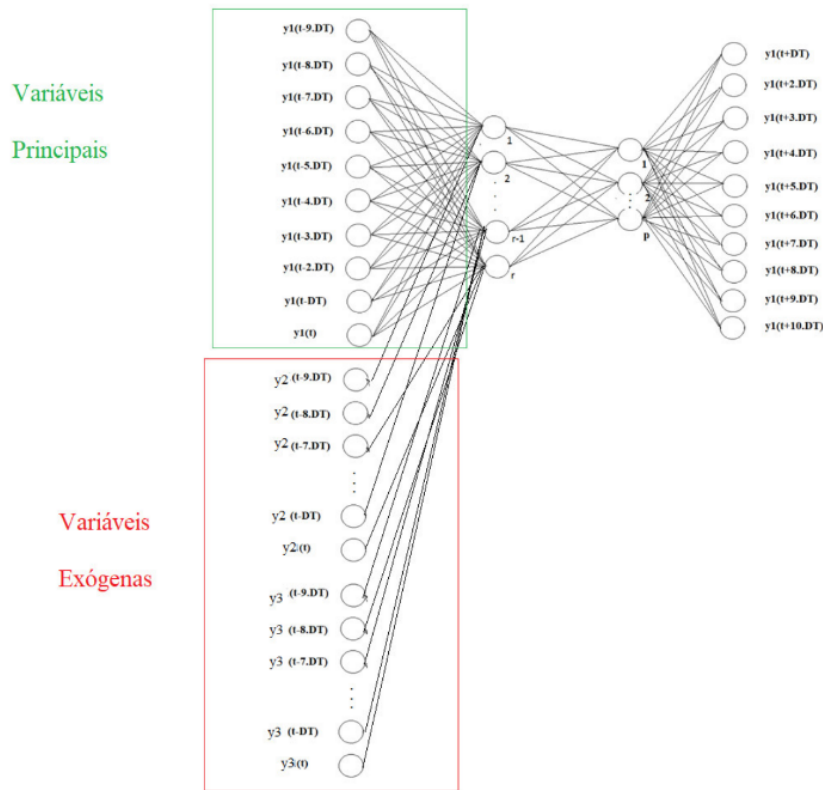


Figura 3: Arquitetura da rede MLP que foi utilizada para resolver o problema de previsão da bolsa de valores com entrada vetorial e saída vetorial.  $r$  é o número total de neurônios na primeira camada interna e  $p$  é o número total de neurônios na segunda camada interna.

Caso nº	Camadas internas	Algoritmo de treinamento empregado
1	r=15 (tansig)	Levenberg-Marquardt (trainlm)
2	r=10 (RELU) e p=10 (RELU)	Levenberg-Marquardt (trainlm)
3	r=p=20 (tansig)	Resilient Backpropagation (trainrp)
4	r=25 (RELU)	Polak-Ribière Conjugate Gradient (traincgp)
5	r=p=25 (tansig)	Bayesian Regularization (trainbr)
6	r=30 (tansig) e p=20 (RELU)	Resilient Backpropagation (trainrp)

Tabela 1: Abordagens utilizadas

## 4 Resultados / Análise

### 4.1 Expoente de Hurst

Para o cálculo do expoente de Hurst das séries temporais, usou-se a função *compute\_Hc* do pacote *hurst*, de Python. Com isso, obteve-se os resultados dispostos nas Figuras 4, 5 e 6.

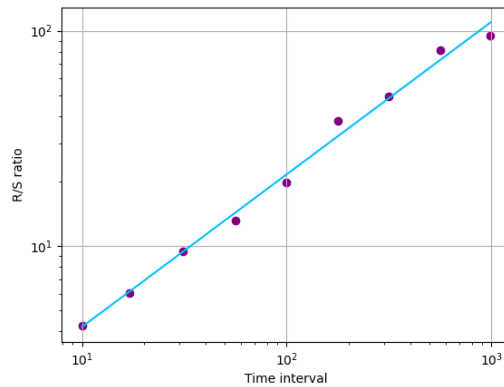


Figura 4: Razão R/S em escala logarítmica para a série da Embraer

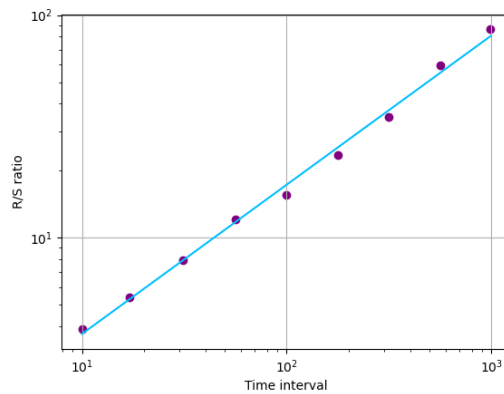


Figura 5: Razão R/S em escala logarítmica para a série da Petrobras

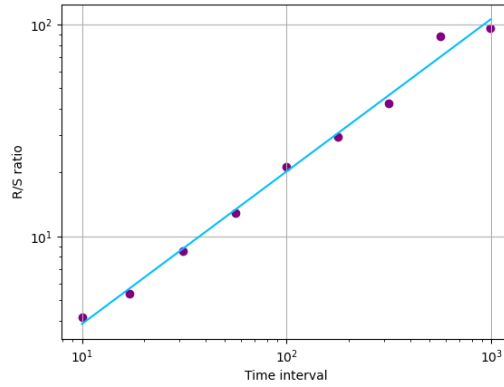


Figura 6: Razão R/S em escala logarítmica para a série da Vale

Assim, obteve-se os valores do expoente de Hurst para cada série, dispostos na Tabela 2.

	Embraer	Petra	Vale
$H$	0.7095	0.6706	0.7201
$c$	0.8198	0.7888	0.7360

Tabela 2: Expoente de Hurst para cada série temporal

Com esses dados, nota-se que as séries temporais analisadas são persistentes ( $0.5 < H < 1$ ), ou seja, são séries que um valor crescente na série tende a ser seguido por outro valor crescente, apresentando um comportamento de tendência.

## 4.2 Previsões na Bolsa de Valores

Os resultados obtidos em cada caso são ilustrados nas Figuras 7, 8, 9, 10, 11 e 12.

O caso 1 exhibe uma sequência de previsões relativamente precisa, mas com ocorrências periódicas de ruídos significativos. Isso pode ser atribuído à escolha da função de ativação tangente sigmoideal (tansig), que pode ser sensível a outliers, resultando em oscilações nos resultados. O uso do Levenberg-Marquardt (trainlm) contribui para a rápida convergência, mas pode ser afetado por ruídos nos dados de treinamento.

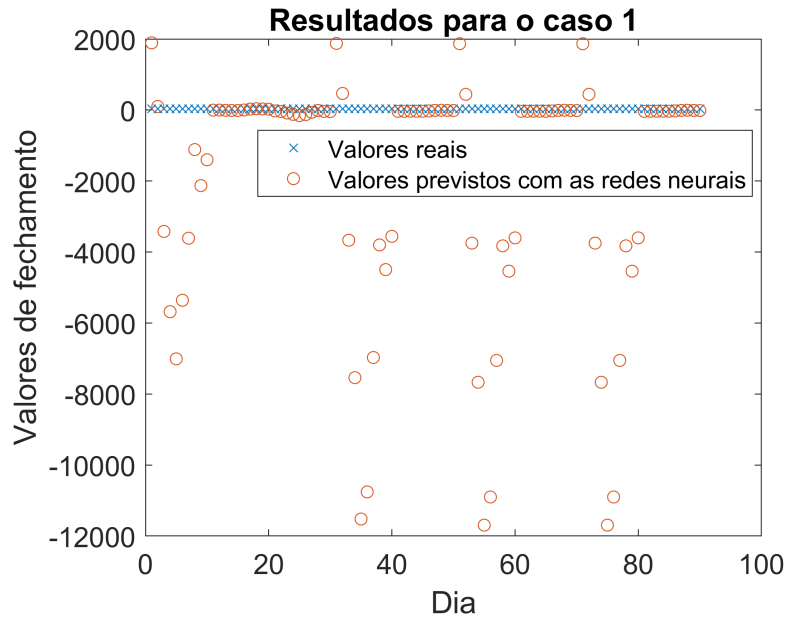


Figura 7: Resultados para o caso 1.

No caso 2, a aplicação da função de ativação Rectified Linear Unit (RELU) nas duas camadas internas proporciona uma tendência mais constante de divergência em relação aos valores reais. A não-linearidade introduzida pela RELU pode levar a uma rápida saturação, resultando em uma perda da capacidade de aprendizado. O Levenberg-Marquardt (trainlm) destaca-se pela convergência eficiente, mas a escolha da função de ativação pode limitar a precisão preditiva.

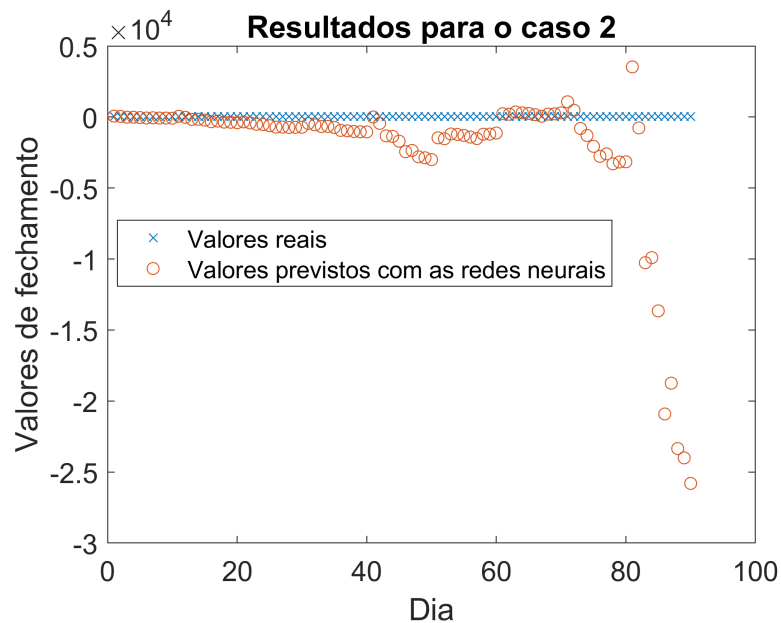


Figura 8: Resultados para o caso 2.

Os resultados do caso 3 apresentam-se próximos da ordem de grandeza dos valores reais, sugerindo uma boa adaptação do modelo. A escolha de 20 neurônios em ambas as camadas internas com a função de ativação tangente sigmoidal (tansig) contribui para uma representação

mais flexível dos padrões subjacentes nos dados. O algoritmo de treinamento Resilient Back-propagation (trainrp) lida bem com ruídos, possibilitando previsões mais robustas.

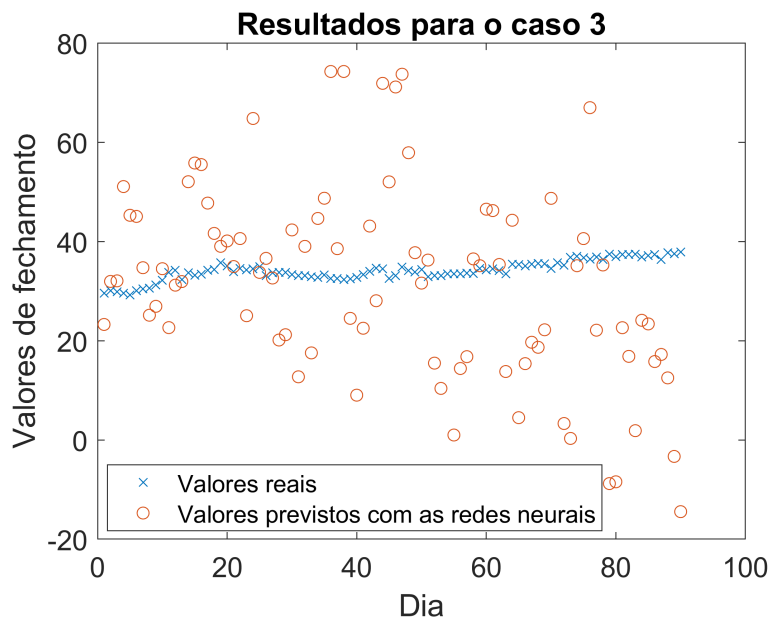


Figura 9: Resultados para o caso 3.

O caso 4 destaca-se pelo maior erro médio quadrático, resultado de uma oscilação final pronunciada nos valores das previsões. A função de ativação RELU, por vezes, pode intensificar a sensibilidade a gradientes, levando a oscilações indesejadas. O Polak-Ribière Conjugate Gradient (traincgp) pode ser menos eficaz em lidar com essas oscilações, contribuindo para uma maior dispersão nos resultados.

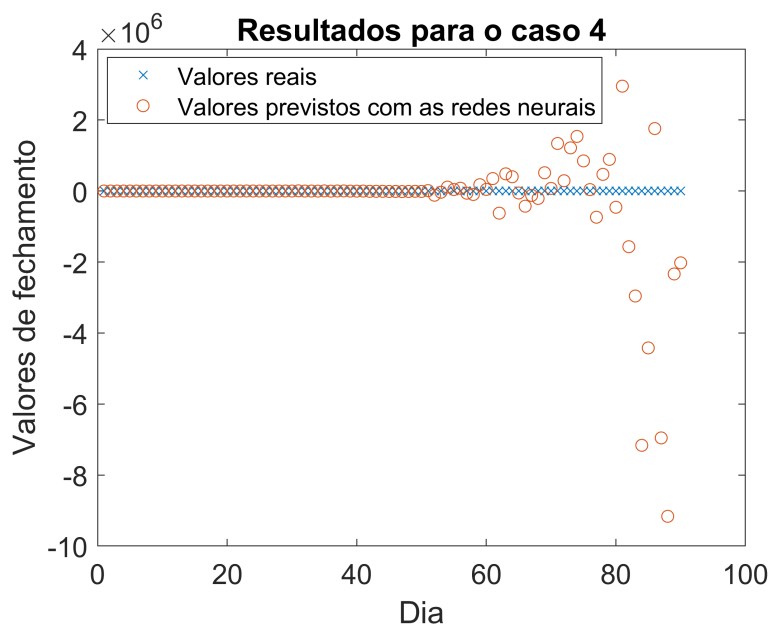


Figura 10: Resultados para o caso 4.

Os resultados do caso 5 revelam uma aproximação aceitável aos valores reais, com uma sequência de previsões dentro de uma faixa aceitável. A escolha de 25 neurônios em ambas



as camadas internas com a função de ativação tangente sigmoidal (tansig) sugere uma boa adaptação do modelo. O algoritmo de treinamento Bayesian Regularization (trainbr) destaca-se por lidar bem com overfitting, contribuindo para uma generalização eficaz.

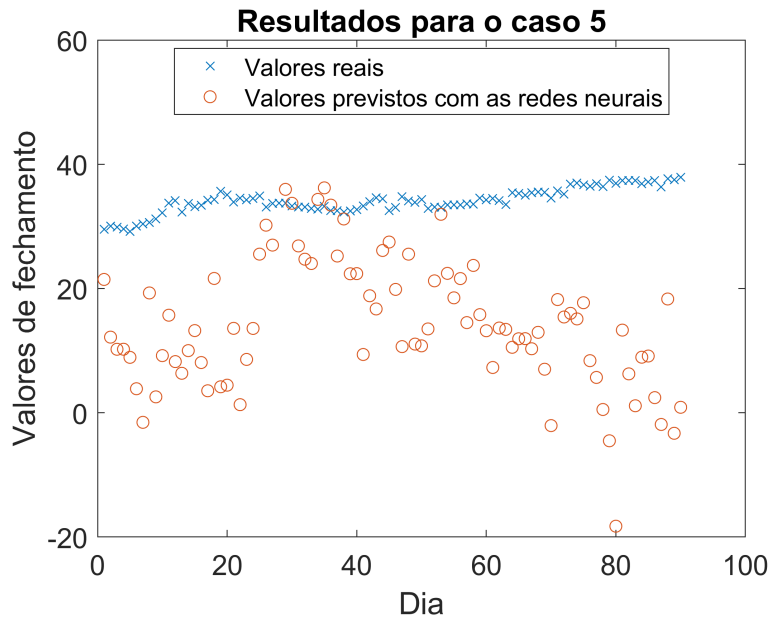


Figura 11: Resultados para o caso 5.

O caso 6 apresenta resultados sequenciais próximos à ordem de grandeza dos valores reais. A combinação de 30 neurônios na primeira camada interna com tansig e 20 neurônios na segunda camada interna com RELU proporciona uma representação flexível dos padrões complexos nos dados. O uso do Resilient Backpropagation (trainrp) contribui para a estabilidade e adaptação eficaz do modelo.

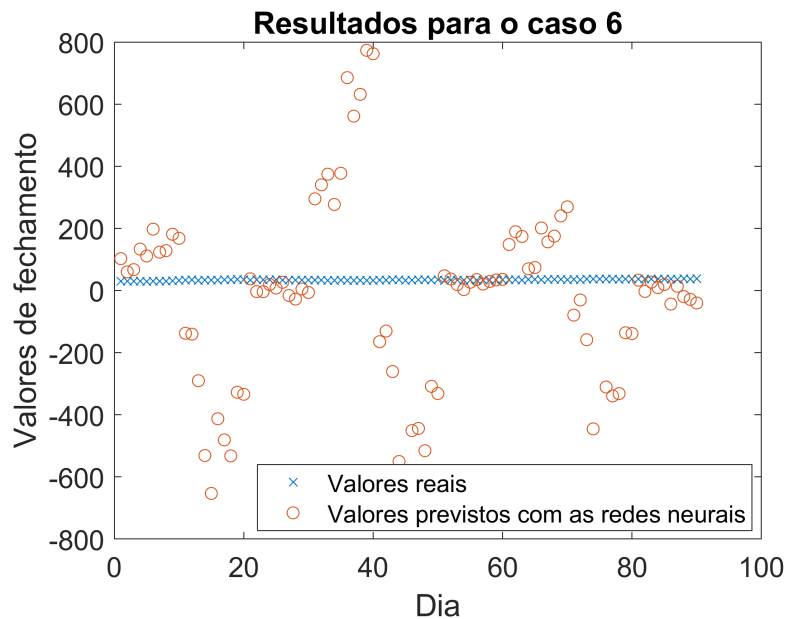


Figura 12: Resultados para o caso 6.

## 5 Conclusão

Em suma, os resultados dos diferentes casos de redes neurais MLP para a previsão de ações da Petrobras sugere que a escolha de parâmetros cruciais, como funções de ativação, número de neurônios e algoritmos de treinamento, desempenha um papel significativo no desempenho preditivo. Os casos 3, 5 e 6 se destacaram, exibindo resultados mais próximos da ordem de grandeza dos valores reais, com menores erros quadráticos médios. Contudo, cada configuração apresentou nuances específicas, ressaltando a necessidade de uma abordagem personalizada na configuração da rede neural para otimizar a precisão das previsões. O entendimento detalhado das características dos dados e a seleção cuidadosa dos parâmetros são fundamentais para o sucesso de modelos de redes neurais aplicados a previsões financeiras.

## 6 Apêndice

### Código utilizado

CT-221 