

Redes Neurais Auto-Regressivas

Thiago Moraes Rizzieri[†]
Orientador(a): José Silvio Govone

Resumo: Neste artigo faremos um breve estudo sobre as Redes Neurais Auto-Regressivas para análise de séries temporais. Buscaremos entender como funciona esta arquitetura particular da rede neural e seu funcionamento para previsões. No fim, faremos uma comparação do desempenho das previsões estimadas entre o modelo de redes neurais e o modelo ARIMA, que é bem conhecido para esta finalidade, com dados de ações da Google de 2015. Vale ressaltar, que meu conhecimento sobre Redes Neurais ainda é bem pouco, por isso, o foco maior do trabalho residirá em seu uso estatístico.

Palavras-chave: Séries temporais; Previsão; Estatística.

1 Redes Neurais Artificiais: O que são?

A rede neural artificial é um modelo matemático que pode apresentar diversas finalidades para o seu uso. Cada finalidade estará associada à uma arquitetura diferente, ou seja, à um formato no qual os neurônios estarão conectados (Figura 10.1). Aqui, iremos analisar a arquitetura multilayer feed-forward para fins de previsão de séries temporais.

A rede é caracterizada pela variedade de “neurônios” presentes nela, que são entradas que recebem valores. No modelo multilayer feed-forward há sempre a presença de uma única camada de entrada e uma única camada de saída, que podem haver um ou mais neurônios em cada. No nosso caso, estaremos interessados em redes com apenas um neurônio na camada da saída.

(espaço em branco intencional)

[†]Bolsista PET/FNDE

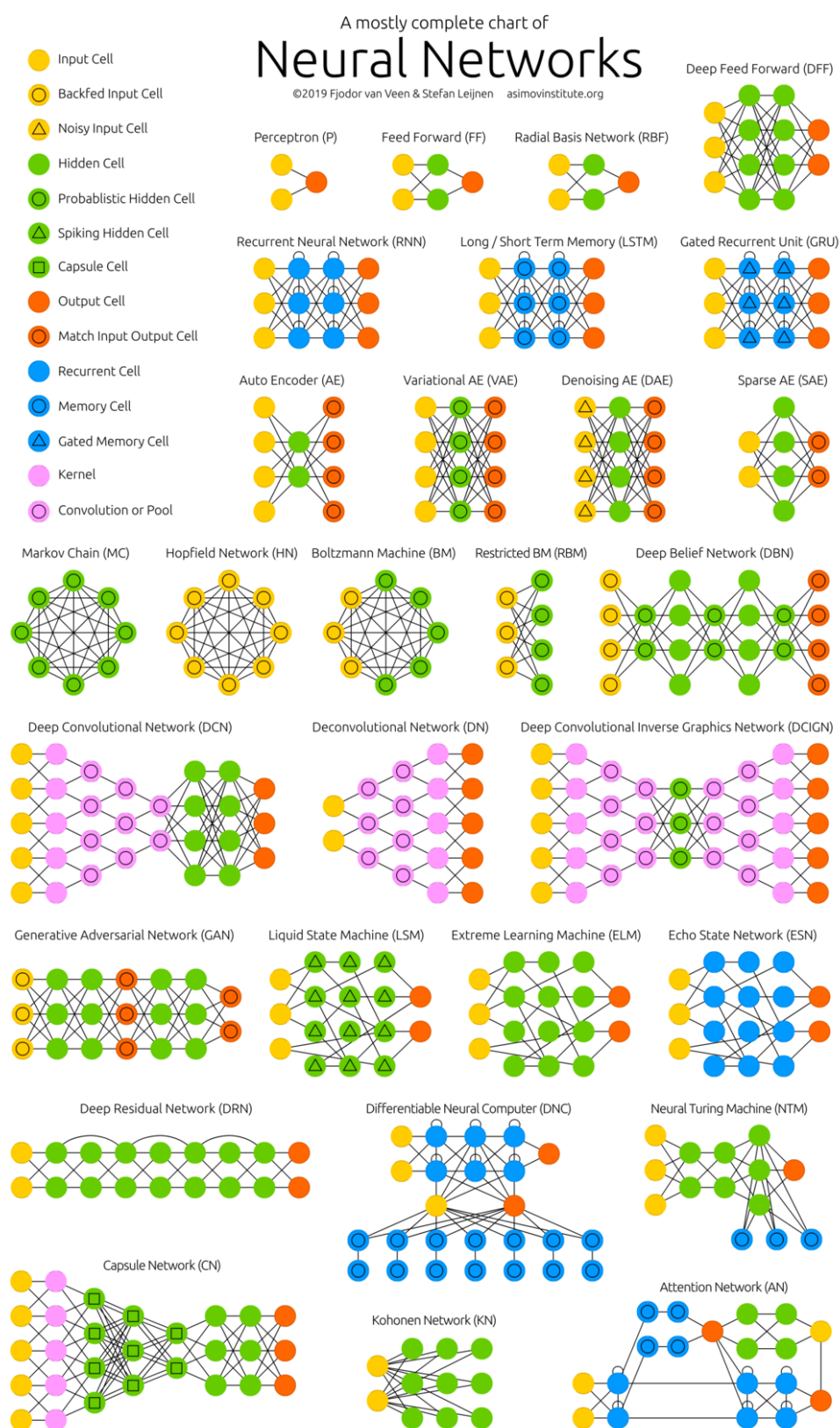


Figura 10.1: Algumas arquiteturas de redes neurais. Fonte: Van Veen, F. & Leijnen, S., The Neural Network Zoo, 2019 [4].

Além dos neurônios, também temos a presença das sinapses, que são as linhas que conectam os neurônios de diferentes camadas. Quando há um certo valor de um neurônio e ele passa para um outro neurônio da próxima camada por uma sinapse, este valor pode ser modificado, seja ele multiplicado por um certo peso (ou parâmetro) ou então sendo transformado por algum tipo de função.

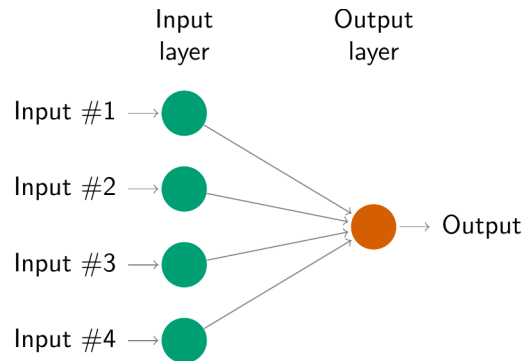


Figura 10.2: Uma arquitetura de rede neural que equivale à um modelo de regressão linear. Fonte: Hyndman, R.J., & Athanasopoulos, G., *Forecasting: principles and practice* [1].

No caso mais simples, havendo apenas uma camada de entrada e uma camada de saída que possui um único neurônio, podemos ter um modelo de regressão linear (Figura 10.2). Temos variáveis na entrada que são multiplicados por certos pesos oferecidos por cada sinapse, no qual é realizado um somatório desses valores todos, juntamente com uma constante adicional. Ou seja, a previsão recebe uma combinação linear das entradas.

Os pesos são selecionados através de uma “função de custo”, como por exemplo, através da média do quadrado dos erros. Assim, os pesos são “aprendidos” (ou estimados) pelos dados. Eles recebem valores aleatórios à princípio, mas são modificados para tornarem a previsão com um menor erro possível.

Quando adicionamos uma camada intermediária com os denominados neurônios ocultos, teremos uma rede neural não linear (Figura 10.3).

Nessa arquitetura, os valores da entrada são multiplicados por alguns pesos e formam uma combinação linear para a camada oculta.

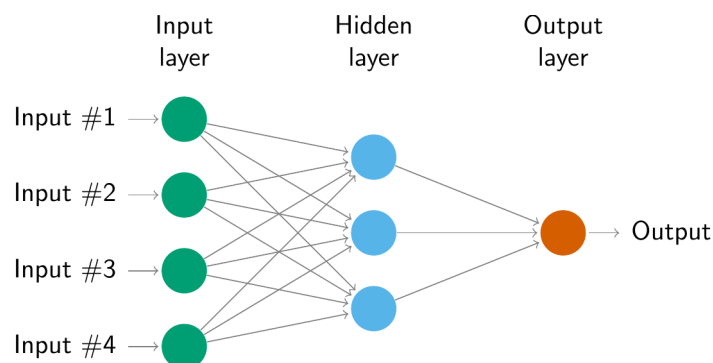


Figura 10.3: Rede Neural com uma camada oculta contendo três neurônios. Fonte: Hyndman, R.J., & Athanasopoulos, G., *Forecasting: principles and practice* [1].

Temos então, no contexto da Figura 10.3, para cada entrada i e cada neurônio oculto j :

$$z_j = \sum_{i=1}^4 w_{i,j} x_i + b_j$$

em que z_j é o valor do neurônio oculto, $w_{i,j}$ é o peso de cada sinapse, x_i é o valor de cada entrada e b_j é a constante de cada neurônio oculto.

Em seguida, da camada oculta para a camada de saída pode ser realizado alguma transformação não linear, como por exemplo, o uso da função logística que é bem comum de ser utilizada.

$$s(z) = \frac{1}{1 + e^{-z}}$$

Essas transformações também são conhecidas como funções de ativação, e são importantes para ajudar se o neurônio deve acolher as informações ou não. Isso pode reduzir o impacto de valores muito diferentes, tornando o modelo mais robusto à valores extremos. Neste artigo, estaremos vendo apenas modelos com uma única camada oculta.

Um estudo mais aprofundado pode ser encontrado na referência [6].

2 Redes Neurais Auto-Regressivas

No modelo de Redes Neurais Auto-Regressivas, os valores de entrada são valores defasados da série temporal, por isso o nome “auto-regressivo”, pois é realizado uma regressão utilizando valores anteriores da própria variável.

Costuma-se utilizar a notação $NNAR(p, k)$ onde p representa a quantidade de entradas e k a quantidade de neurônios ocultos.

Pode ser interessante dar um olhar com maior atenção para dados que apresentam sazonalidade, ou seja, quando há a presença de períodos específicos que sofrem naturalmente uma certa mudança esperada. Como por exemplo, a venda de sorvetes que é afetada pelas estações do ano, pois vende-se mais quando está quente, ou seja, no verão. Nestes casos, ainda é interessante adicionar os últimos valores sazonais para realizar previsões, seja qual for o período de sazonalidade. Dessa forma, denotamos por $NNAR(p, P, k)m$ onde P representa quantas entradas de valores sazonais estaremos utilizando e m representa o período sazonal.

Por padrão, a escolha das p entradas é realizado pelo teste de Akaike igual é feito para os modelos auto-regressivos $AR(p)$. Caso tenha sazonalidade, o valor P por padrão é 1, ou seja, por padrão se utiliza um único valor sazonal. A quantidade de neurônios ocultos k é determinado por $k = (p + P + 1)/2$ arredondado para o maior número inteiro mais próximo.

Como o modelo não é bem definido, é difícil estabelecer um intervalo de previsão da forma usual. Por isso, seu intervalo é formado através de uma quantia grande de simulações, que por padrão são feitas 1000 vezes.

3 O Modelo ARIMA

O modelo ARIMA é um modelo bem comum e conhecido ao se estudar séries temporais, combinando o modelo auto-regressivo (AR), que compõe valores passados dos dados, o modelo de médias móveis (MA), que compõe os erros passados dos dados, e a diferenciação para tornar a série estacionária.

Diferenciar um dado, equivale a fazer a diferença de seu valor com o valor anterior. Assim, diferenciar y_t equivale à $y'_t = y_t - y_{t-1}$.

Uma série é estacionária quando suas propriedades estatísticas não se alteram com o tempo, ou seja, a sua média e variância são constantes ao longo de todo o tempo. Além disso, sua covariância dependerá apenas do intervalo entre os dois valores utilizados.

Assim, temos o modelo $ARIMA(p, q, d)$ onde p representa a quantidade de valores passados utilizados, d a quantidade de erros passados utilizados e q a quantidade de vezes que a série precisa ser diferenciada para ser estacionária.

4 Metodologia

Comparamos os modelos para o estudo das ações da Google nos dias úteis de 2015, em dólares. Analisamos o desempenho dos modelos fornecendo apenas os dados de 2015 e fazendo-os prever em relação à janeiro de 2016.

Os dados são utilizados através do comando `gafa_stock` que possui dados dos preços de ações da Google, da Amazon, do Facebook e da Apple, no período de 2014 à 2018. Coletamos os dados pelo *Yahoo Finance historical data* [5].

Realizamos o ajuste de modelo pelo programa R, assim como todos os processos e cálculos.

5 Comparação entre os modelos

Automaticamente, o programa R nos ajustou um modelo $ARIMA(0, 1, 1)$ e um modelo $NNAR(2, 2)$ para os dados. Podemos observar a previsão de cada um dos modelos nas Figuras 10.4 e 10.5. Temos também a comparação dos modelos $ARIMA$ e $NNAR$ com o que realmente aconteceu em janeiro de 2016 (Figura 10.6).

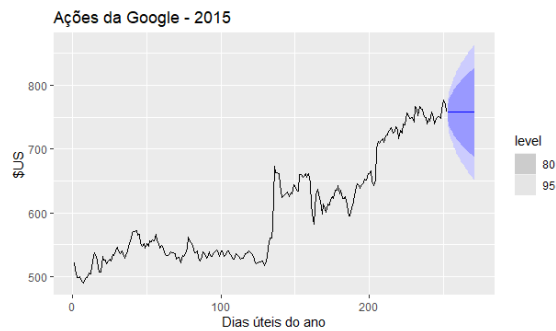


Figura 10.4: Modelo $ARIMA$ com seu intervalo de previsão. Fonte: elaborado pelo autor.

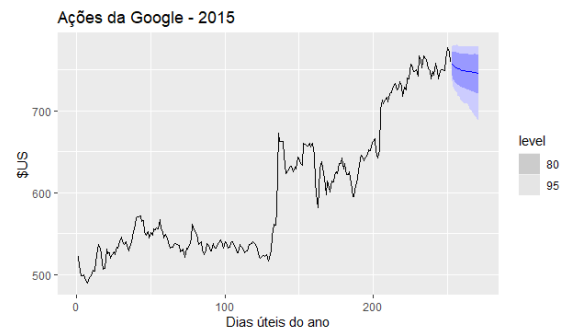


Figura 10.5: Modelo $NNAR$ com seu intervalo de previsão. Fonte: elaborado pelo autor.

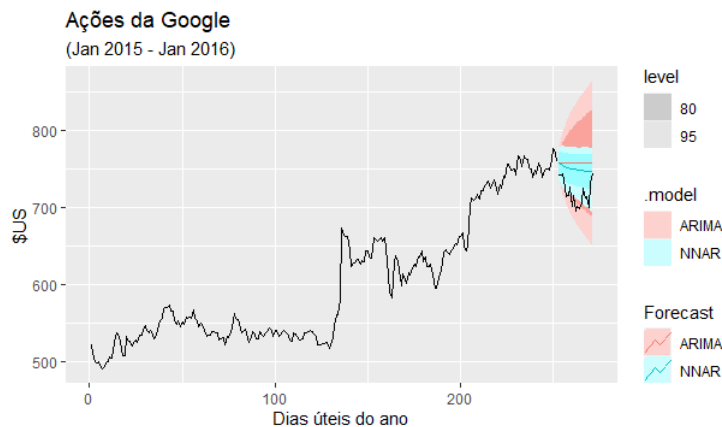


Figura 10.6: Modelos $ARIMA$ e $NNAR$ sobrepostos na previsão de Janeiro de 2016 juntamente com os dados originais deste período. Fonte: elaborado pelo autor.

Logo com estes gráficos já podemos observar que as amplitudes do intervalo de confiança do modelo NNAR é menor do que as do modelo ARIMA, indicando um melhor desempenho do modelo NNAR. Porém, apenas esta análise não é suficiente para garantir que o modelo NNAR se ajustou melhor aos dados do que o modelo ARIMA.

Avaliamos o gráfico dos resíduos de cada modelo (Figura 10.7 e Figura 10.8). Em ambos os modelos, observamos uma forte aleatoriedade e baixa autocorrelação entre erros passados. Assim, consideramos que os erros assumem o formato de um ruído branco.

Além disso, notamos que há uma baixa variabilidade até o momento de pico. Isso se deve à mudança brusca da estrutura da série temporal, que se encontrava estável. Depois disso, os resíduos voltaram a se estabilizar, mas não tanto quanto antes.

Vale ressaltar também, que grande parte dos resíduos são próximos ou iguais ao valor zero, indicando que o modelo é bem fiel ao que realmente acontece.

Para avaliar a precisão das previsões pontuais, levamos em conta três métodos comuns que nos fornecem um bom resumo do comportamento dos resíduos. Todos fazem uso do valor de erro e_t entre as previsões dos modelos com o que realmente ocorreu em janeiro de 2016. Na referência [1] podemos encontrar mais detalhes sobre estes métodos.

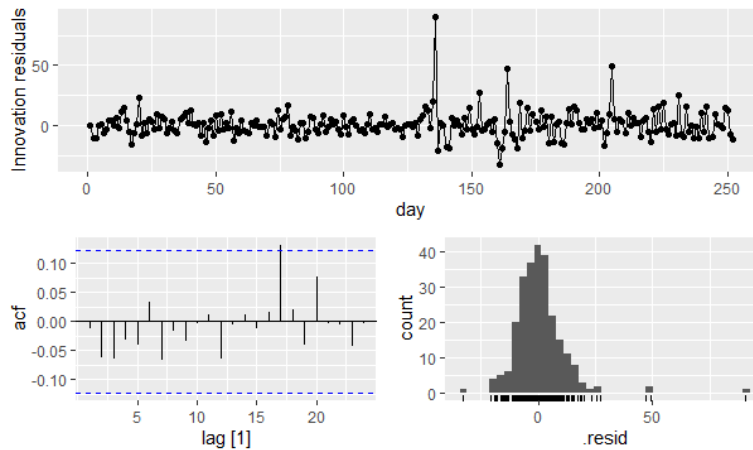


Figura 10.7: Gráfico da série temporal, do correlograma e do histograma dos resíduos do modelo ARIMA. Fonte: elaborado pelo autor.

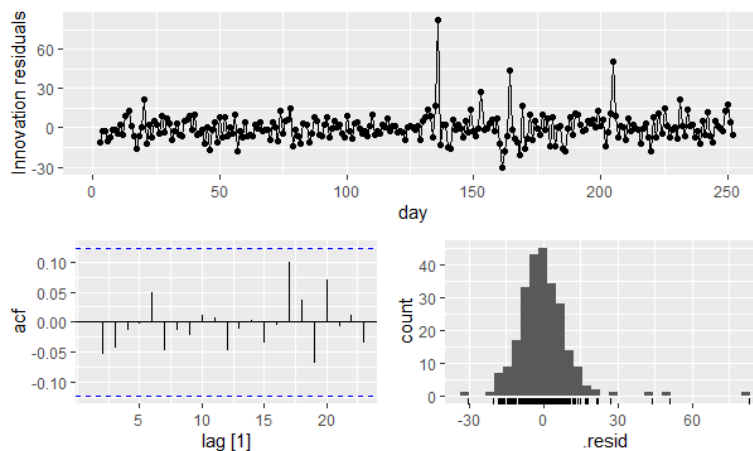


Figura 10.8: Gráfico da série temporal, do correlograma e do histograma dos resíduos do modelo NNAR. Fonte: elaborado pelo autor.

Temos o MAE que representa a média do erro absoluto, o RMSE que representa a raiz do erro quadrático médio e MAPE que representa a média do erro percentual absoluto, dados respectivamente por:

$$\text{MAE} = \text{média}(|e_t|) \quad \text{RMSE} = \sqrt{\text{média}(e_t^2)} \quad \text{MAPE} = \text{média}(|p_t|)$$

em que $p_t = 100e_t/y_t\%$.

Os modelos tiveram os seguintes resultados para estes valores.

Método	MAE	RMSE	MAPE
ARIMA	39.0	42.2	5.49
NNAR	31.8	35.1	4.48

Tabela 10.1: Tabela da precisão da previsão pontual dos modelos

Observamos que os valores do MAE, MAPE e RMSE são menores para o modelo NNAR. Concluímos com base nos resultados acima que o modelo NNAR se ajustou melhor para estes dados.

Agradecimentos: Primeiramente, gostaria de agradecer ao Programa de Educação Tutorial, PET, pelo apoio financeiro e pela jornada de orientações referentes à iniciação científica. Agradeço também a amigos e família pelo apoio diário, que me mantiveram firme nos estudos.

Abstract: In this article, we will make a brief study on Autoregressive Neural Networks for time series analysis. We will seek to understand how this particular neural network architecture works and how it works for predictions. In the end, we will make a comparison of the performance of estimated predictions between the neural network model and the ARIMA model, which is well known for this purpose, with Google stock data from 2015. It is noteworthy, that my knowledge of Neural Networks is still very little, so the main focus of the work will reside in its statistical use.

Keywords: Time series; Forecasting; Statistics.

Referências Bibliográficas

- [1] Hyndman, R.J., Athanasopoulos, G., *Forecasting: principles and practice*, 3ª edição, OTexts: Melbourne, Australia. Disponível em: <https://otexts.com/fpp3>. Acesso em: 26 de outubro de 2021.
- [2] Reza, Aashiq., *An Approach to Make Comparison of ARIMA and NNAR Models For Forecasting Price of Commodities*. Towards Data Science, 2020. Disponível em: <https://towardsdatascience.com/an-approach-to-make-comparison-of-arima-and-nnar-models-for-forecasting-price-of-commodities-f80491aeb400>. Acesso em: 26 de outubro de 2021.
- [3] Data Science Academy, *Deep Learning Book*, 2021. Disponível em: <https://www.deeplearningbook.com.br>. Acesso em: 26 de outubro de 2021.
- [4] Van Veen, F. & Leijnen, S., *The Neural Network Zoo*, The Asimov Institute, 2019. Disponível em: <https://www.asimovinstitute.org/neural-network-zoo>. Acesso em: 26 de outubro de 2021.
- [5] Yahoo! Finance. Disponível em: <https://finance.yahoo.com>. Acesso em: 26 de outubro de 2021.

- [6] Hastie T., Tibshirani R., Friedman J., *The Elements of Statistical Learning Data Mining, Inference and Prediction*, 2^a edição, Springer, 2009.