

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**“PREVISÃO DO ÍNDICE BURSÁTIL IBEX 35 USANDO
REDES NEURAIS ARTIFICIAIS”**

Salvador Falcón Canillas

Orientador: Prof. Dr. Carlos Roberto Minussi

Dissertação de mestrado apresentada à Faculdade de Engenharia - UNESP – Câmpus de Ilha Solteira, como requisito para obtenção do título de **Mestre em Engenharia Elétrica**.

Área de concentração: Automação.

Ilha Solteira – SP, abril-2021

FICHA CATALOGRÁFICA

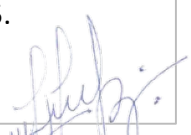
Desenvolvido pelo Serviço Técnico de Biblioteca e Documentação

F181p Falcón Canillas, Salvador.
Previsão do índice bursatil IBEX 35 usando redes neurais artificiais /
Salvador Falcón Canillas. -- Ilha Solteira: [s.n.], 2021
72 f. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de
Engenharia de Ilha Solteira. Área de conhecimento: Automação, 2021

Orientador: Carlos Roberto Minussi
Inclui bibliografia

1. Previsão. 2. Ibex 35. 3. Bolsa de valores. 4. *Backpropagation*. 5.
Feedforward. 6. Regressão generalizada.



João Josué Barbosa
Serviço Técnico de Biblioteca e Documentação
Diretor Técnico
CRB 8-5642

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Previsão do Índice Bursatil IBEX 35 por Redes Neurais Artificiais

AUTOR: SALVADOR FALCÓN CANILLAS

ORIENTADOR: CARLOS ROBERTO MINUSSI

Aprovado como parte das exigências para obtenção do Título de Mestre em ENGENHARIA ELÉTRICA, área: Automação pela Comissão Examinadora:

Prof. Dr. CARLOS ROBERTO MINUSSI (Participação Virtual)

Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira - UNESP

Profª. Drª. ANNA DIVA PLASENCIA LOTUFO (Participação Virtual)

Departamento de Engenharia Elétrica / Faculdade de Engenharia de Ilha Solteira - UNESP

Profa. Dra. THAYS APARECIDA DE ABREU SANTOS (Participação Virtual)

Centro Estadual de Educação Tecnológica "Paula Souza" - Faculdade de Tecnologia de Americana

Ilha Solteira, 09 de abril de 2021

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Carlos Roberto Minussi, por depositar em mim a confiança para realizar esta pesquisa, como seu orientado, ajudando-me tanto com as dificuldades próprias do mestrado como com outras dificuldades pessoais.

A todos os professores que contribuíram para o desenvolvimento desta tese, em especial:

Prof^a. Dr^a. Anna Diva Plasencia Lotufo, Prof^a. Dr^a. Mara Lúcia Martins Lopes é Prof^a. Dr^a.Thays Aparecida de Abreu Santos.

A Haislán Bernardes, pelo seu apoio desde o início, ajudando-me com todos os problemas pessoais e burocráticos necessários, porque sem a sua ajuda a conclusão desta tese teria sido impossível.

A Mauro de Souza Tonelli Neto, pela sua grande ajuda na abordagem da tese, tanto técnica como pessoal.

A Giovanni Panegossi Formaggio por ser um bom amigo e um excelente companheiro, que esteve sempre disponível para me ajudar com questões da tese e que foi para mim um guia importante especialmente no início da dissertação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

RESUMO

A previsão de índices bolsistas de diferentes bolsas de valores é uma das questões mais importantes para economistas e investidores, a fim de conhecer, antecipadamente, os movimentos que ocorrem no mercado de investimento. Os investidores confiam, frequentemente, no seu conhecimento da economia dos mercados, dos fatores macroeconômicos dos países e de outros recursos, tais como: o comportamento dos gráficos de ações, para fazerem os seus investimentos da forma mais eficiente possível. O fato de serem pessoas que usam seu julgamento para investir no mercado de ações torna-se uma tarefa complicada, se não impossível, avaliar todos os fatores envolvidos no mercado de ações. Além dessa incapacidade, há fatores como elementos psicológicos e julgamentos errôneos subjetivos que podem levar o investidor a ter perdas econômicas, ou a não aproveitar boas oportunidades de investimento. Por essas razões, ao longo das últimas décadas, foram feitas tentativas de prever índices bolsistas, com diferentes sistemas entre eles:

Modelos lineares, tais como ARIMA, Garch e Arch, redes neurais tais como Feedforward Neural Network (FNN), Generalized Regression Neural Network (GRNN), Probabilistic Neural Network (PNN), etc. e a combinação de modelos lineares e redes neurais.

Nesta pesquisa propõe-se uma forma como realizar a previsão do valor no dia seguinte do índice de ações espanhol IBEX 35, utilizando três redes neurais artificiais diferentes e comparando o desempenho de cada uma dessas redes em diferentes situações. O IBEX 35 é o principal índice de referência do mercado acionário da bolsa espanhola. É formado pelas 35 empresas mais líquidas listadas no Sistema de Interconexão da Bolsa de Valores Espanhola nas quatro bolsas de valores espanholas. As redes neurais a serem empregadas são *Feedforward*, *Cascade-Forward* e *Generalized Regression*

Palavras-chave: Previsão. Ibex 35. Bolsa de valores. *Backpropagation*. *Feedforward*. *Cascade*. Regressão generalizada.

ABSTRACT

Forecasting stock market indices from different stock exchanges are one of the most important issues for economists and investors, to know in advance, the movements that occur in the investment market. Investors often rely on their knowledge of the economy of the markets, the macroeconomic factors of the countries, and other resources, such as the behaviour of stock charts, to make their investments as efficiently as possible. Being people who use their judgment to invest in the stock market is complicated, if not impossible, to evaluate all the factors involved in the stock market. In addition to this inability, there are factors such as psychological elements and subjective misjudgements that may lead the investor to have economic losses or not take advantage of good investment opportunities. For this reason, over the last decades, attempts have been made to predict stock market indices, with different systems among them:

Linear models, such as ARIMA, Garch, and Arch, neural networks such as the Feedforward Neural Network (FNN), Generalized Regression Neural Network (GRNN), Probabilistic Neural Network (PNN), etc. and the combination of linear models and neural networks.

This research presents how to forecast the next day value of the Spanish IBEX 35 stock index using three different artificial neural networks and comparing the performance of each of these nets in different situations. The IBEX 35 is the main reference index of the Spanish stock market. It is formed by the 35 most liquid companies listed on the Spanish Stock Exchange's Interconnection System on the four Spanish stock exchanges. The neural networks to use are Feedforward, Cascade-Forward, and Generalized Regression Neural Network.

Keywords: Forecast. Ibex 35. Stock market. Backpropagation. Feedforward. Cascade. Generalized regression.

SUMÁRIO

1	INTRODUÇÃO.....	8
2	ESTADO DA ARTE	11
3	REDES NEURAIIS UTILIZADAS.....	22
3.1	<i>FEEDFORWARD NEURAL NETWORK(FNN)</i>	22
3.2	<i>CASCADE-FORWARD NEURAL NETWORK(CFNN)</i>	32
3.3	<i>GENERALIZED REGRESSION NEURAL NETWORK(GRNN)</i>	33
4	METODOLOGIAS	39
4.1	OBTENÇÃO DAS ENTRADAS	40
4.2	PROCESAMENTO DAS ENTRADAS	41
5	RESULTADOS OBTIDOS.....	43
5.1	ANÁLISE DO MAPE DO CASO 1	44
5.1.1	USANDO A REDE NEURAL <i>FEEDFORWARD</i>	45
5.1.2	USANDO A REDE NEURAL <i>CASCADE-FORWARD</i>	48
5.1.3	USANDO A <i>GENERALIZED REGRESION NEURAL NETWORK</i>	51
5.2	ANÁLISE DO MAPE DO CASO 2.....	54
5.2.1	USANDO A REDE NEURAL <i>FEEDFORWARD</i>	55
5.2.2	USANDO A REDE NEURAL <i>CASCADE-FORWARD</i>	58
5.2.3	USANDO A <i>GENERALIZED REGRESION NEURAL NETWORK</i>	61
6	CONCLUSÕES	64
	REFERÊNCIAS	66

1 INTRODUÇÃO

Hoje em dia, a previsão dos índices bolsistas das diferentes bolsas internacionais é uma questão importante na economia global, uma vez que movimenta enorme quantidade de dinheiro. Muitos fatores influenciam se esses índices podem cair ou subir de valor.

Os fatores que influenciam estes índices podem ser provenientes de índices macroeconômicos nacionais, tais como: dívida externa, produção industrial, desemprego, prêmio de risco, reputação do banco nacional, taxa de juros, etc. (MACIÁN-PÉREZ, 2017)

Outros fatores, que podem influenciar os índices bolsistas de um país, são os índices bolsistas de outros países, o Brexit (a saída do Reino Unido da União Europeia), o preço das matérias-primas (ferro, ouro, petróleo etc.), guerras econômicas, conflitos militares, epidemias etc.

Como mencionado anteriormente, a previsão de índices bolsistas é uma tarefa muito importante do ponto de vista econômico e a sua previsão é crucial quando se fazem investimentos no mercado financeiro. Embora haja muitos economistas e especialistas em investimentos que confiam na sua própria experiência e conhecimentos para investir, eles são limitados pelas próprias condições humanas como: fadiga, erros subjetivos e a incapacidade de analisar enormes quantidades de dados.

Portanto, a previsão de índices bolsistas utilizando redes neurais artificiais (RNA) (HAYKIN, 2008) parece ser uma ideia bastante interessante para realizar esta tarefa, pois se tem um longo histórico de dados sobre os preços dos índices bolsistas diários de todo o mundo pode-se obter previsões de excelente precisão. Estes dados podem ser encontrados em páginas, tais como: *Yahoo Finance* ou *Investing*.

Há muitos artigos onde são utilizados diferentes métodos para prever os índices bolsistas de diferentes países, dos quais o artigo "PREDICCIÓN DEL ÍNDICE BURSÁTIL IBEX 35 MEDIANTE SELECCIÓN DE INPUTS CON REDES NEURONALES" (FERNÁNDEZ-BENLLOCH, 2016) foi aquele em que se obteve a maior inspiração para este trabalho. Nele, é feita uma previsão do índice IBEX 35 utilizando a rede neural *Feedforward* e usando dados dos seguintes índices como entrada:

- Futuro IBEX
- IBEX
- DAX
- Futuro de Petróleo
- STANDARD AND POORS
- FTSE
- DOW JONES Industrial
- DOW JONES Transportes
- EUR/USD

Nesta pesquisa também foi realizada a previsão do índice IBEX 35, utilizando a técnica de janelamento. É realizada a previsão empregando-se três redes neurais diferentes, comparando o desempenho de cada uma delas em dois casos diferentes. No primeiro caso usou-se como entrada só os dados do mesmo IBEX 35 é no segundo caso foram usados outros índices bolsistas, commodities¹ e taxas de câmbio tomadas como entrada que seriam os seguintes:

- AEX (Holanda)
- BEL 20 (Bélgica)
- CAC 40 (França)
- DAX 30 (Alemanha)
- HANG SENG (Hong Kong)
- NIKKEI 225 (Japão)
- OMXS 30 (Suécia)
- S&P ASX 200 (Austrália)
- S&P TSX 60 (Canadá)
- SMI 30 (Suíça)
- Taiwan Weighted (Taiwan)

- IBEX 35 (Espanha)
- Milho
- Ouro
- Prata
- USD/JPY (taxa de câmbio entre o dólar americano e o iene japonês)
- EUR/CAD (taxa de câmbio entre o euro e o dólar canadense)
- EUR/AUD (taxa de câmbio entre o euro e o dólar australiano)
- EUR/JPY (taxa de câmbio entre o euro e o iene japonês)

As redes neurais a serem utilizadas são *Feedforward* (HAYKIN, 2008), *Cascade-Forward* (WARSITO, 2018) e Regressão Generalizada (*GRNN*) (SPECHT, 1991).).

- 1: São aqueles produtos destinados ao uso comercial, e que têm como característica mais relevante, que não têm nenhum valor agregado, não são processados, ou não têm nenhuma característica diferenciadora em relação aos outros produtos que se encontram no mercado. Por este motivo são utilizados como matéria-prima para fabricar outras matérias elaboradas.

2 ESTADO DA ARTE

Economistas de todo o mundo sempre estudaram e tentaram compreender as flutuações de preços na bolsa de valores, porque os investimentos na bolsa de valores estão sujeitos a riscos, os retornos são variáveis e a sua existência é incerta. A previsão da bolsa de valores é um tema de interesse, particularmente para aqueles que investem na bolsa de valores. Por conseguinte, a previsão de um índice bolsista é muito desafiante e foram pesquisados vários métodos para prever o mais eficientemente possível.

Os modelos lineares, em razão à fácil interpretação dos seus elementos, possuem uma vantagem considerável sobre os modelos mais complexos, o que levou a sua utilização em muitas aplicações; uma delas tem sido a previsão de séries temporais financeiras. Nos anos 80, os modelos utilizados, para modelar a média de um bem, eram o *Arima* ou *Arimax*, mas nos anos 90 os modelos *ARCH* (ENGLE, 1982). e *Garch* (BOLLERSLEV, 1986) dão maior importância à volatilidade do que a média dos dados.

Em Franses e Ghijssels (1999) confirmaram que os modelos de *Garch* são úteis para prever a volatilidade do mercado bolsista, mas os resíduos estimados destes modelos apresentam, frequentemente, a demasiada curtose e não captam todas as características dos dados.

Chang, Yeung e Yip (2000) indicaram a influência de indicadores econômicos, tais como: a taxa de emprego, ou o índice de preços do consumidor, pois cada um deles é uma medida em relação a um detalhe da economia, pelo que relacionaram cada indicador por meio de um gráfico ao índice da bolsa de valores para compreender até que ponto as variáveis relacionadas com a bolsa de valores eram bem conhecidas.

Em 2003, os estudos continuam: Lasfer, Melnik & Thomas documentam o comportamento dos preços de ações, a curto prazo, após um período de tensão na bolsa, concentrando-se no comportamento diário dos preços da bolsa via 39 índices da bolsa. A investigação mostrou que os desempenhos anormais após o stress são significativamente mais elevados para os mercados emergentes e a dimensão dos saltos após o período de stress está relacionada com a liquidez do mercado. (LASFER; MELNIK; THOMAS, 2003)

Bhardwaj e Swanson (2006) sugeriram um novo modelo chamado *Arfima*, que faz estimativas usando uma variedade de procedimentos padrão que fornecem

previsões significativamente melhores do que *AR*, *MA*, *ARMA*, *Garch* e modelos relacionados, com base na análise dos erros médios quadráticos da previsão (*MSFE*) e na utilização de testes de precisão preditiva.

Liu e Hung (2010) realizaram um estudo no qual, utilizando a volatilidade diária do índice americano S&P-100, compararam os modelos *Garch-N*, *Garch-t*, *Garch-HT* e *Garch-SGT* com modelos do tipo assimétrico, tais como *GJR-Garch* e *Egarch*. Os resultados obtidos mostram que estes últimos melhoram a volatilidade da previsão e demonstram que o componente assimétrico é mais importante do que a especificação da distribuição quando estão presentes assimetria, leptocurtose e efeitos de Alavancagem.

Contudo, a economia destaca-se pela sua instabilidade e pela presença de não linearidade nos seus ciclos, o que torna os modelos lineares não inteiramente eficientes. Por esta razão, a exploração de dados é atualmente caracterizada pela utilização crescente de modelos de previsão não lineares.

Komo, Chang e Ko (1994) utilizam modelos estatísticos não paramétricos e não lineares porque muitas relações importantes na área das finanças têm este tipo de relação. As redes neurais artificiais têm a propriedade de capturar as características não lineares dos índices da bolsa e mostraram que podem ser treinadas com uma quantidade suficiente de informação para identificar tais relações não lineares entre os valores de entrada e saída. A aplicação de redes neurais à previsão de séries financeiras cresceu nos últimos anos, uma vez que parece ser um método eficaz e apresenta uma multiplicidade de oportunidades.

Zemke (1999) afirma que os mercados, com um menor volume de operações, são mais fáceis de prever. Utilizou a Bolsa de Varsóvia como referência e propôs as técnicas de *Machine Learning (ML)* (HAYKIN, 2008) para a sua análise. A tarefa é prever o valor do índice via decisões binárias, ou seja, prever se o valor do índice WIG na semana de negociação está acima ou abaixo do valor atual. As quatro técnicas *ML* utilizadas são: previsão de rede neural, classificador Bayesiano (STIGLER, 1983), *K-nearest neighbor* (SILVERMAN, 1989) e *K-nearest neighbor prediction scrutinized* (Previsão do Vizinho mais próximo) (AL-DOSARY, 2019). Com estas técnicas concluiu-se que utilizando métodos *K-nearest neighbor*, com 64% de eficácia e redes neurais, se obtêm melhores previsões porque o mercado bolsista é dominado pela não linearidade dos dados. (ZEMKE, 1999).

No mesmo ano, Lee e Jo desenvolveram um intérprete gráfico para prever o mercado, a que chamaram *Candlestick Chart Analysis Expert System* (Sistema Especialista em Análise Gráfica de Velas) (MORRIS, 2006). Este sistema possui padrões e regras que podem prever os movimentos futuros dos preços das ações. Os padrões definidos são classificados em cinco grupos no que diz respeito ao seu significado: queda, subida, neutralidade, tendência contínua, e padrões em que a tendência é invertida. Os resultados experimentais obtidos por Lee e Jo revelaram que o modelo teve uma taxa média de sucesso de 72%, o que ajuda os investidores a obterem maiores benefícios do seu investimento em ações. (LEE; JO, 1999).

O modelo de Chen, Leung e Daouk (2003) prevê a mudança no índice da Bolsa de Valores de Taiwan por meio de outro modelo de rede neural artificial chamado Rede Neural Probabilística (*PNN*); este método utiliza dados históricos do índice da Bolsa de Valores de Taiwan, e os resultados mostram que as estratégias de investimento baseadas nele obtêm retornos mais elevados do que outras estratégias, tais como métodos generalizados de momentos (*GMM*). A superioridade deste modelo deve-se à sua capacidade de identificar inconsistências e desinformação.

Toro, Molina e Garcés (2006) apresentaram um estudo comparativo entre a previsão de preços na bolsa de valores utilizando redes neurais e neuro-nebulosas (*neuro-fuzzy*). É apresentada uma metodologia aplicável à previsão do comportamento de qualquer tipo de ações, com base no cálculo preliminar da correlação entre o preço e outras variáveis de mercado. A previsão do valor das ações em bolsa afeta os processos de tomada de decisão das empresas ou outros agentes do mercado afetados pelo seu comportamento. Os resultados obtidos foram satisfatórios em termos da percentagem de erro do valor estimado em relação ao valor real. As técnicas inteligentes provam ser uma ferramenta interessante que está mostrando resultados promissores e a abrir um leque de possibilidades a explorar no campo da previsão.

Chaigusin, Chirathamjaree e Clayden (2008) mostram um caso particular em que a previsão do índice da Bolsa de Valores da Tailândia é realizada por meio de redes neurais do tipo *feedforward*. Os dados de entrada utilizados correspondem ao índice SET, Dow Jones, Nikkei, Hang Seng, o preço do ouro, a taxa mínima de empréstimo e as taxas de câmbio entre o thai-baht tailandês e do dólar, mostrando via pesquisa que os

movimentos do índice da Bolsa de Valores da Tailândia são sensíveis às entradas acima referidas.

Alonso e García (2009) utilizam dados de alta frequência para encontrar um modelo estatístico que reduza o grau de incerteza quanto ao comportamento do futuro imediato (próximos dez minutos) dos operadores na bolsa colombiana, especificamente do IGBC; para tal, utilizam modelos de *Garch* em média (*Garch-M*), que levam em conta a alavancagem, o dia da semana, a hora e os efeitos dia-hora.

Em Lu, Chang, Chen, Chiu e Lee (2009), a previsão do índice da Bolsa de Xangai B-Share é utilizada para comparar as metodologias *MARS* (Splines de Regressão Adaptativa Multivariada), *BPNN* (Rede Neural de Retropropagação), *SVM* (Máquina de Vetores de Suporte) (CORTES; VAPNIK, 1995.) e *MLR* (Regressão Linear Múltipla) (XIN, 2009). Os resultados obtidos mostraram que o modelo *MARS* fornece melhores previsões em termos de erro e de precisão do que os outros, com uma taxa de erro de 1,15% e uma precisão direcional (*DA*) de 82%.

Guresen, Kayakutlu e Daim (2011) observaram que na maioria dos casos, os modelos de redes neurais permitem obter melhores resultados do que outros métodos. Os dados utilizados correspondem aos valores diários do índice Nasdaq de 7 de Outubro de 2008 a 26 de Junho de 2009. Os modelos que foram comparados neste trabalho são o Perceptron Multicamada (*MLP*) (ROSENBLATT, 1958), Rede Neural Artificial Dinâmica (*DAN2*) (VELÁSQUEZ-HENAO, 2012) e um modelo híbrido de redes neurais e *Garch*. O desempenho dos modelos foi avaliado usando-se o Erro Quadrático Médio (*MSE*) e o Desvio Médio Absoluto (*MAD*), obtendo-se que o modelo clássico *MLP* excede os outros, obtendo melhores resultados.

Outra forma de fazer uma previsão seria utilizando um sistema híbrido, ou seja, combinando modelos lineares como os já apresentados anteriormente com redes neurais. Na previsão da bolsa de valores muitos autores fizeram combinações de métodos ou modelos de previsão procurando incorporar as vantagens de cada um dos modelos anteriores no método híbrido.

Kim e Han (2000) expressam que existe uma grande quantidade de estudos que incluem modelos de Redes Neurais Artificiais (RNAs). No entanto, as RNAs têm limitações na aprendizagem de padrões por causa do fato de os dados do mercado bolsista terem uma dimensionalidade complexa e um grande ruído. Usando o algoritmo genético

(GA) é possível selecionar o tipo de rede neural de tal forma que a característica relevante do subconjunto possa ser otimizada. GA concentra-se, neste trabalho, na discretização dos dados, simplificando assim o processo de aprendizagem para as RNAs e reduzindo o ruído e os dados redundantes.

Pai e Lin (2005) propõem uma metodologia híbrida entre *ARIMA* e *SVM*. Este último é uma nova técnica de rede neural que tem sido aplicada com sucesso na resolução de problemas de estimativa de regressão não linear. Esta metodologia é aplicada a problemas de previsão de preços de ações, alcançando resultados promissores. O modelo apresentado melhora muito o desempenho da previsão do preço das ações do modelo *ARIMA* e do modelo *SVM* individualmente, uma vez que tanto em teoria como empiricamente, a hibridação de dois modelos diferentes reduz os erros de previsão.

Parisi, Parisi e Diaz (2006) utilizaram modelos dinâmicos multivariados construídos a partir de algoritmos genéticos e modelos de redes neurais para prever o sinal de variações semanais nos índices bolsistas asiáticos. O sinal da projeção foi comparado com a variação observada em cada *i*-ésima semana, obtendo-se que os modelos multivariados com algoritmos genéticos são mais robustos e permitem obter melhores índices de rentabilidade do que os modelos construídos com redes neurais, pelo que por meio desta técnica o investidor pode tomar posições em ativos mais voláteis, comprando-os quando se encontram na parte inferior do ciclo básico de evolução de preços e vendendo-os quando se encontram na parte superior.

Roh (2007) realizou um estudo no qual utiliza o teste *ADF* (Dickey-Fuller Aumentado) para verificar a estabilidade e o teste *ARCH LM* para verificar a heterocedastidade dos dados do índice Kospi 200 da Coreia. Depois, é feita uma comparação do poder preditivo de diferentes modelos individuais tais como *EWNA*, *Garch*, *Egarch* e RNAs, e modelos híbridos, obtendo como resultado que o erro médio absoluto (*MSE*) foi menor para os híbridos *NN-Garch* e *NN-Egarch*, o que nos permite concluir que os modelos de RNAs juntamente com modelos de séries temporais podem melhorar o poder preditivo em termos de direção e de precisão.

Yudong e Lenan (2009), por outro lado, propõem que se as redes neurais *feedforward* usando o algoritmo *backpropagation* tenham sido utilizadas para prever índices, neste trabalho, são incorporadas com Optimização da Quimiotaxia Bacteriana Melhorada (*IBCO*) (MULLER, 2002), os dados utilizados são os do índice S&P 500. O

Erro Quadrático Médio (*MSE*) diminui consideravelmente, e o tempo de treinamento da rede foi menor do que o modelo *BP* individualmente. O modelo *IBCO-BP* oferece menos complexidade computacional, maior precisão na previsão e menos tempo de treino.

El-Henawy *et al.* (2010) usam redes neurais Perceptron Multicamadas (*MLP*) para prever o índice Kospi 200 (índice de preços de ações da Coreia) ao longo de um período de onze anos. Para melhorar a sua arquitetura, parâmetros, e aumentar a precisão das previsões e diminuir o tempo de treino, são utilizados três algoritmos: Recozimento Simulado mais conhecido como *Simulated Annealing* (*SA*) (KIRKPATRICK, 1983), Algoritmo Genético (*GA*) e um híbrido que combina os dois anteriores. Os resultados mostram que o melhor algoritmo é o *SA*, que excede a abordagem híbrida em 30% de precisão, e 40% ao algoritmo genético. *SA* requer 7 minutos de treinamento, *GA* 73 minutos e a abordagem híbrida 98 minutos.

Boyacioglu e Avci (2010) confirmam que os retornos do mercado de ações podem ser previstos via desempenho passado e de variáveis macroeconômicas e financeiras. Prever os retornos das ações é uma tarefa complexa devido ao grande número de fatores envolvidos no mercado de ações. Isto torna as séries de preços de ações dinâmicas, não lineares, complicadas e caóticas. Este artigo utiliza o modelo *ANFIS*, que combina a teoria da rede neural e a lógica *fuzzy* utilizando variáveis que compreendem outros índices bolsistas e variáveis macroeconômicas, o índice de preços no consumidor, a produção industrial, a taxa de câmbio do dólar, etc. O modelo obtido tem um R-quadrado do 98%, o que significa que a *ANFIS* prevê com um elevado grau de certeza o índice da Bolsa de Valores de Istambul.

Wang, Cheng e Hsu (2010) incorporam a cadeia Markov no modelo estocástico difuso e utilizam dados do índice da Bolsa de Valores de Taiwan. Este modelo é capaz de considerar, simultaneamente as taxas de câmbio, o aumento das probabilidades e as quedas dos índices bolsistas. Dos 330 ensaios realizados durante a experiência de três meses, o modelo foi significativamente melhor em 298 deles.

Yu e Huarng (2010) procuraram aplicar redes neuronais para implementar um novo modelo de série temporal difusa que melhora as previsões e inclui os vários graus de pertença no estabelecimento de relações difusas que ajudam na captura das relações mais apropriadas. Estas relações difusas são utilizadas para prever o índice da Bolsa de Valores de Taiwan.

De acordo com Cheng, Chen e Wei (2010), foram encontrados dois inconvenientes em muitos dos modelos de previsão do passado: 1) os pressupostos estatísticos sobre variáveis que são necessárias para modelos de séries temporais, tais como: *ARMA* e *ARCH* para modelos de previsão de equações matemáticas, e estes não são facilmente compreendidos pelos investidores em ações; 2) as regras extraídas de alguns algoritmos de inteligência artificial (*IA*), tais como redes neurais, não são facilmente realizadas.

Para superar estes inconvenientes, é proposto um modelo de previsão híbrido, utilizando indicadores multitécnicos para prever as tendências dos preços das ações. Além disso, inclui quatro procedimentos propostos no modelo híbrido a fim de fornecer regras eficazes de previsão, que foi desenvolvido a partir das regras extraídas da teoria de conjuntos aproximados (*RST*): 1) selecionar indicadores técnicos chave; 2) utilizar a Abordagem de Distribuição de Probabilidade Cumulativa (*CPDA*) (CHUNG-HO, 2010).; 3) utilizar um algoritmo *RST* (LUPU, 2008) para extrair regras linguísticas do conjunto de dados de indicadores técnicos linguísticos; 4) utilizar algoritmos genéticos (*GA*) para refinar as regras extraídas para uma melhor precisão de previsão. A eficácia do modelo proposto é verificada por dois tipos de avaliações de desempenho: precisão e valores de retorno, e utilizando um período de seis anos do índice Taix de Taiwan como conjunto de dados de teste. Os resultados experimentais mostram que o modelo proposto é superior aos modelos de previsão, a teoria do conjunto aproximado (*RST*) e o algoritmo genético em termos de precisão, e as avaliações de desempenho da bolsa revelaram que os benefícios produzidos pelo modelo proposto são superiores aos modelos comprar e manter (*buy and hold*), *RST* e *GAS*. (CHENG; CHEN; WEI, 2010).

Shen *et al.* (2011) selecionaram uma rede neural do tipo Função de Base Radial (*RBF*) para prever os índices bolsistas da Bolsa de Xangai. A fim de otimizar a rede e aumentar a eficiência das previsões, é utilizado o Algoritmo de Enxame de Peixes Artificial (*AFSA*). Os resultados do *RBF* otimizado com *AFSA*, Algoritmo Genético (*GA*) (HOLLAND, 1992) e Otimização por enxame de partículas (*PSO*) (KENNEDY, 1995) são tão bons como as previsões feitas com *ARIMA*, *Backpropagation* (*BP*) ou Máquina de Vetores de Suporte (*SVM*), apresentando o *RBF* com *AFSA* não a mais alta precisão, mas sendo um novo algoritmo inteligente é bem-sucedido em aumentar o resultado das previsões da rede original da Função de Base Radial (*RBF*). (SHEN *et al.*, 2011).

Wang *et al.* (2011) explicam que, devido à variedade de fatores que afetam o mercado de ações, é proposta a utilização do algoritmo *Wavelet De-noising Backpropagation (WDBP)*; os dados utilizados correspondem aos do índice da Bolsa de Valores de Xangai de Janeiro de 1993 a Dezembro de 2009. O modelo *WDBP* foi comparado com um modelo *Backpropagation*, resultando o *WDBP* superior na precisão das previsões.

A previsão da bolsa de valores é uma das questões mais importantes no campo financeiro. Portanto, propõem um modelo chamado Algoritmo Genético/ Otimização por Enxame de Partículas (*GA/PSO*), que será avaliado usando-se índices como o Nasdaq100, Dow Jones e o S & P500. A rede neural utilizada como preditor é a Rede Neural Recorrente de Sigmoide Diagonal (*SDRNN*), porque a sua arquitetura diminui o erro e aumenta a precisão em muitas aplicações. (ABOUELDAHAB; FAKHRELDIN, 2011).

Asadi *et al.* (2012) propõem um modelo híbrido que é uma combinação de métodos de pré-processamento de dados, algoritmos genéticos e algoritmos Levenberg-Marquardt (LM) (LEVENBERG, 1944), utilizados para a aprendizagem da alimentação de redes neurais na previsão da taxa de ações. Também utilizam informação de métodos de pré-processamento, tais como transformação de dados e seleção de variáveis de entrada para melhorar a precisão do modelo. A competência do método proposto é testada com a sua aplicação para prever alguns índices de ações utilizados na literatura. Os resultados mostram que o método proposto é capaz de lidar com as flutuações nos valores das existências e também obtém uma boa precisão preditiva. Portanto, pode ser utilizado para modelar relações complexas entre entradas e saídas ou para encontrar padrões nos dados enquanto se faz a previsão financeira.

Wang *et al.* (2012) apresentam o modelo híbrido (*PHM*), que é uma combinação do *ARIMA*, Modelo de Alisamento Exponencial (*ESM*) e uma Rede Neural *Feedforward (FNN)* usando o algoritmo *backpropagation*, que permite capturar características lineares e não lineares ao longo de uma série temporal. Para testar o modelo, foram utilizados dados mensais do índice SZII da China e o índice Dow Jones *Industrial Average Index (DJIAI-USA)*. A Previsão Direcional (*AD*) é utilizada para avaliar a exatidão das previsões. Os resultados obtidos mostram que o modelo *PHM*

híbrido proporciona melhores resultados em termos de erro e precisão do que outros modelos como *ARIMA*, *ESM* e *FNN* por causa da sua robustez.

Dai, Wu e Lu (2012) propõem um modelo de previsão de séries cronológicas combinando análises não lineares de componentes independentes e redes neurais para prever os índices bolsistas asiáticos. Exemplos representativos são o Nikkei 225 e o Shanghai B Share. Os resultados experimentais mostram que o modelo de previsão proposto não só melhora a precisão da previsão da abordagem da rede neural, mas também excede os três métodos de comparação com precisões superiores a 80%.

Muitos analistas acreditam que os movimentos que ocorrem na bolsa de valores estão intimamente ligados ao comportamento da economia do país onde opera, pelo que existem algumas variáveis que podem influenciar a bolsa de valores pública: crescimento econômico, taxas de juros, o preço do dólar, produção nacional, preços do petróleo, comportamento da inflação, desemprego, entre outras variáveis da macroeconomia de um país.

Pierdzioch, Dopke e Hartmann (2008) comparam previsões de mercados bolsistas voláteis com base em dados macroeconômicos em tempo real e dados macroeconômicos revistos. Para tal são utilizados dois critérios estatísticos, um baseado na utilidade e o outro nas opções. O principal resultado é que o valor estatístico e econômico das previsões baseadas em dados macroeconômicos em tempo real é comparável ao valor estatístico e econômico das previsões baseadas em dados macroeconômicos revistos.

Cai, Chou e Li (2009) publicam um estudo no qual investigam as correlações dinâmicas entre alguns índices bolsistas internacionais e a inflação. Utilizam um modelo de série temporal autorregressivo chamado (*DSTCC-CARR*), que revela como as correlações são significativamente variáveis ao longo do tempo e como estão relacionadas com flutuações cíclicas nas taxas de inflação e volatilidade do mercado. Os índices internacionais utilizados são os da França, Alemanha, Rússia, Hong Kong, Japão e Estados Unidos, e os níveis de correlação foram significativos em 95%. As correlações mais elevadas ocorrem quando os países se encontram numa fase de contração e as mais baixas quando se encontram numa fase de expansão.

Chen (2009) sugere que as variáveis macroeconômicas servem para prever recessões no mercado de ações dos EUA (o índice de preços *Standard & Poor's* (S & P

500)) (referência), especialmente os diferenciais de rendimento das curvas e as taxas de inflação são os indicadores mais úteis; isto foi concluído após a utilização de métodos paramétricos e não paramétricos para identificar períodos de recessão no mercado de ações e considerando tanto dentro como fora da amostra, testes de capacidade de previsão das variáveis.

Majhi *et al.* (2009) propõem a previsão de índices bolsistas utilizando a Optimização da Forragem Bacteriana Adaptativa (*ABFO*) e Optimização da Forragem Bacteriana (*BFO*), a estrutura utilizada nestes modelos de previsão é um simples combinador linear. Os novos modelos são computacionalmente mais eficientes, com uma previsão mais precisa, e mostram uma convergência mais rápida em comparação com outros modelos computacionais evolutivos, tais como o algoritmo genético.

Gençtürk, Çelik e Binici (2012) descobriram que não só a relação entre as variáveis macroeconômicas e a bolsa de valores é importante, mas também que as ligações causais entre as relações de curto e longo prazo são importantes. As variáveis macroeconômicas tais como a taxa de juros, a taxa de câmbio e o índice de preços no consumidor afetam as decisões dos investidores. Foi utilizado um modelo *VECM*, com dados da Bolsa de Valores turca, com o qual se obteve que não existe uma relação a longo prazo entre o índice da Bolsa de Valores turca e o índice de preços ao consumidor, a taxa de juros e a taxa de câmbio do dólar, mas existe entre o índice e a produção industrial, o que leva a concluir que a produção industrial é um forte fator de previsão da Bolsa de Valores turca.

Caldas e Pires (2012) forneceram provas empíricas sobre a influência das variáveis macroeconômicas e do risco-país no principal índice da bolsa de valores Brasileiro (IBOVESPA). As provas práticas foram obtidas com a aplicação de Quadrados Mínimos Ordinários (*OLS*), o Método de Generalização de Momentos (*GMM*) e sistemas *GMM*. Os resultados mostram que a política monetária e a gestão da dívida pública, bem como a credibilidade e a reputação, afetam o risco país e o desempenho do mercado bolsista brasileiro.

Como conclusão podemos dizer que existem muitos modelos e métodos que têm sido utilizados nos últimos tempos para prever índices bolsistas em países de todo o mundo. Os primeiros modelos eram autorregressivos, conhecidos como *ARIMA*, que simplesmente captavam as características lineares do mercado bolsista e a sua vantagem

reside no facto de não exigirem diferentes séries de dados, implicando uma poupança na especificação e identificação do modelo no sentido da econometria tradicional, mas devido ao grande número de relações não lineares entre os dados foi necessário implementar modelos não lineares e não paramétricos.

Depois, a fim de melhorar a precisão das previsões, foram introduzidos modelos híbridos que agrupavam as vantagens dos dois ou mais modelos envolvidos. Em particular, a descoberta da influência de variáveis macroeconômicas em algumas bolsas de valores do mundo impulsionou a concepção de modelos que utilizam tais variáveis como dados de entrada para prever índices bolsistas, provando ser muito eficazes.

Isto não significa que modelos autorregressivos como *ARIMA* ou *Garch* sejam obsoletos, mas sim que devido à sua estrutura não podem captar as características não lineares presentes nas séries cronológicas financeiras, tais como as séries de índices de ações, tornando necessário combinar estes métodos com outros que possam modelar tais características e melhorar as previsões.

Este caminho deixou como resultado uma longa lista de métodos para prever o comportamento da bolsa de valores ou das ações individuais da empresa. É necessário continuar a propor e a adaptar os métodos de previsão às necessidades, aos contextos, às características dos índices bolsistas, de modo a obter previsões mais precisas, minimizando o risco do investidor. (CECILIA-GARCÍA, *et al.*, 2013).

3 REDES NEURAIS UTILIZADAS

3.1 FEEDFORWARD NEURAL NETWORK(FNN)

Estas redes são um tipo de rede de aprendizagem supervisionada e funcionam, basicamente, aplicando um padrão de entrada à rede, que se propaga desde a primeira camada por meio das camadas superiores da rede até gerar uma saída. (GUEVARA-DÍAZ, 2007).

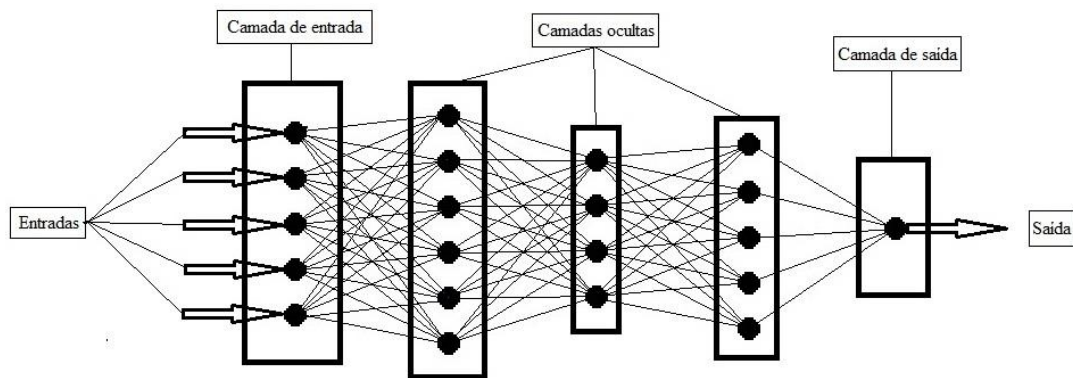


Figura 1- Arquitetura do *Feedforward*

Fonte: Elaboração própria.

A equação matemática da arquitetura da Figura 1 pode ser escrita da seguinte forma:

$$y = f^0\left(\sum_{j=1}^k w_j^o f_j^h\left(\sum_{i=1}^n w_{ji}^h x_i\right)\right) \quad (1)$$

Onde y é a saída, f^0 é a função de activação na camada de saída e f_j^h é uma função de activação na camada oculta e se um viés é adicionado à camada de entrada e a função de activação de cada neurónio na camada oculta é f^h , então a equação (1) torna-se:

$$y = f^0\left(w^b + \sum_{j=1}^k w_j^o f_j^h\left(w_j^b \sum_{i=1}^n w_{ji}^h x_i\right)\right) \quad (2)$$

Onde w^b é o peso do enviesamento à saída e w_j^b é o peso do enviesamento à camada oculta. (WARSITO; SANTOSO, 2018)

Neste caso o tipo de algoritmo de treinamento usado foi o Levenberg-Marquardt *Backpropagation* (LMBP) (MARQUARDT, 1963) que é um dos algoritmos

supervisionados mais rápidos e é altamente recomendado, embora exija mais memória computacional que outros algoritmos.

O algoritmo Levenberg-Marquardt é uma variação do método de Newton que foi concebido para minimizar funções que são somas de quadrados de outras funções não lineares. Isto é bem adequado para o treino de redes neurais, onde o índice de desempenho é o erro quadrático médio (*MSE*). Considerando-se a forma do método de Newton em que o índice de desempenho é uma soma de quadrados. O método de Newton para otimizar um índice de desempenho $F(\mathbf{x})$ é:

$$x_{k+1} = x_k - A_k^{-1} g_k, \quad (3)$$

$$\text{onde } A_k \equiv \nabla^2 F(x) \big|_{x=x_k} \text{ e } g_k \equiv \nabla F(x) \big|_{x=x_k}$$

Assumindo-se que $F(x)$ é uma função de soma de quadrados:

$$F(x) = \sum_{i=1}^N v_i^2(x) = v^T(x)v(x) \quad (4)$$

então o j -ésimo elemento do gradiente seria:

$$[\nabla F(x)]_j = \frac{\partial F(x)}{\partial x_j} = 2 \sum_{i=1}^N v_i(x) \frac{\partial v_i(x)}{\partial x_j} \quad (5)$$

Portanto o gradiente pode ser escrito em forma de matriz:

$$\nabla F(x) = 2 J^T(x) v(x) \quad (6)$$

Onde

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial v_1(\mathbf{x})}{\partial x_1} & \frac{\partial v_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial v_2(\mathbf{x})}{\partial x_1} & \frac{\partial v_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial v_N(\mathbf{x})}{\partial x_1} & \frac{\partial v_N(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial v_N(\mathbf{x})}{\partial x_n} \end{bmatrix} . \quad (7)$$

é a matriz Jacobiana.

A seguir mostra-se a matriz Hessiana. Os elementos j e k da matriz de Hessiana seriam:

$$[\nabla^2 F(x)]_{k,j} = \frac{\partial^2 F(x)}{\partial x_k \partial x_j} = 2 \left\{ \sum_{i=1}^N \frac{\partial v_i(x)}{\partial x_k} \frac{\partial v_i(x)}{\partial x_j} + v_i(x) \frac{\partial^2 v_i(x)}{\partial x_k \partial x_j} \right\} \quad (8)$$

A matriz Hessiana pode então ser expressa em forma de matriz:

$$\nabla^2 F(x) = 2J^T(x)J(x) + 2S(x) \quad (9)$$

Onde

$$S(x) = \sum_{i=1}^N v_i(x) \nabla^2 v_i(x) \quad (10)$$

Considerando-se o valor de $S(x)$ pequeno, podemos aproximar a matriz Hessiana como:

$$\nabla^2 F(x) \cong 2J^T(x)J(x) \quad (11)$$

Então se substitui-se a Equação (11) e Equação (6) na Equação (3), é obtido o método Gauss-Newton:

$$\begin{aligned}
x_{k+1} &= x_k - [2J^T(x_k)J(x_k)]^{-1} 2J^T(x_k)v(x_k) \\
&= x_k - [J^T(x_k)J(x_k)]^{-1} J^T(x_k)v(x_k)
\end{aligned} \tag{12}$$

Pode-se observar que a vantagem de Gauss-Newton em relação ao método padrão de Newton é que não requer o cálculo de segundas derivadas.

Um problema com o método Gauss-Newton é que a matriz $\mathbf{H} = \mathbf{J}^T \mathbf{J}$ não pode ser invertível. Isto pode ser ultrapassado através da seguinte modificação da matriz Hessiana aproximada:

$$\mathbf{G} = \mathbf{H} + \mu \mathbf{I} \tag{13}$$

Para ver como esta matriz pode ser invertida, suponha que os valores próprios e os vectores próprios de $\mathbf{H}$ são $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e $\{z_1, z_2, \dots, z_n\}$. Então:

$$\mathbf{G}z_i = [\mathbf{H} + \mu \mathbf{I}]z_i = \mathbf{H}z_i + \mu z_i = \lambda_i z_i + \mu z_i = (\lambda_i + \mu) z_i \tag{14}$$

Portanto, os autovetores de $\mathbf{G}$ são os mesmos que os autovetores de $\mathbf{H}$, e os autovalores de $\mathbf{G}$ são $\lambda_i + \mu$. $\mathbf{G}$ pode ser definido como positivo aumentando μ até $(\lambda_i + \mu) > 0$ para todos os valores de i e por tanto a matriz será invertível

Isto leva ao algoritmo Levenberg-Marquardt:

$$x_{k+1} = x_k - [J^T(x_k)J(x_k) + \mu_k \mathbf{I}]^{-1} J^T(x_k)v(x_k) \tag{15}$$

ou

$$\Delta x_k = -[J^T(x_k)J(x_k) + \mu_k \mathbf{I}]^{-1} J^T(x_k)v(x_k) \tag{16}$$

Este algoritmo tem a característica muito útil de, à medida que o μ_k é aumentado, aproxima-se do algoritmo de descida mais acentuado com uma pequena taxa de aprendizagem:

$$x_{k+1} \equiv x_k - \frac{1}{\mu_k} J^T(x_k)v(x_k) = x_k - \frac{1}{2\mu_k} \nabla F(x) \tag{17}$$

para uma μ_k grande, enquanto à medida que μ_k é reduzido a zero, o algoritmo se torna Gauss-Newton.

O algoritmo começa com μ_k definido para algum valor pequeno (por exemplo, $\mu_k = 0,01$). Se um passo não render um valor menor para $F(\mathbf{x})$, então o passo é repetido com multiplicado por algum fator $\vartheta > 1$ (por exemplo, $\vartheta = 10$). Eventualmente $F(\mathbf{x})$ deve diminuir, pois estaríamos dando um pequeno passo na direção da descida mais íngreme. Se um passo produzir um valor menor para $F(\mathbf{x})$, então μ_k é dividido por ϑ para o próximo passo, de modo que o algoritmo se aproximará de Gauss-Newton, o que deve proporcionar uma convergência mais rápida. O algoritmo oferece uma boa relação entre a velocidade do método de Newton e a convergência garantida da descida mais íngreme.

A seguir observa-se como aplicar o algoritmo Levenberg-Marquardt ao problema do treinamento em redes multicamadas. Se cada meta ocorrer com igual probabilidade, o erro quadrático médio é proporcional à soma dos erros quadráticos sobre as metas $\mathbf{Q}$ no conjunto de treinamento:

$$\begin{aligned} F(x) &= \sum_{q=1}^Q (t_q - a_q)^T (t_q - a_q) \\ &= \sum_{q=1}^Q e_q^T e_q = \sum_{q=1}^Q \sum_{j=1}^{S^M} (e_{j,q})^2 = \sum_{i=1}^N v_i^2 \quad (18) \end{aligned}$$

onde $e_{j,q}$ é o elemento j do erro para o par q entrada($\mathbf{a}$)/alvo($\mathbf{t}$). Eq. (18) é equivalente ao índice de desempenho, Eq. (4), para o qual Levenberg-Marquardt foi projetado. Portanto, deve ser uma questão simples adaptar o algoritmo para o treinamento da rede. Teoricamente isto é verdade, mas requer algum cuidado na elaboração dos detalhes.

A etapa chave no algoritmo Levenberg-Marquardt é o cálculo da matriz Jacobiana. Para realizar este cálculo utiliza-se uma variação do algoritmo *backpropagation*. Lembrando que no procedimento padrão de *backpropagation* calcula-se as derivadas dos erros quadráticos, com respeito aos pesos e enviesamentos da rede. Para criar a matriz jacobiana, precisamos computar as derivadas dos erros, em vez das derivadas dos erros quadráticos.

É uma questão simples, conceitualmente, modificar o algoritmo de *backpropagation* para calcular os elementos da matriz Jacobiana. Infelizmente, embora o

conceito básico seja simples, os detalhes da implementação podem ser um pouco complicados.

Antes de apresentarmos o procedimento para o cálculo do jacobiano, vamos dar uma olhada de perto em sua forma (Eq. (7)).

Observe-se que o vetor erro é:

$$v^T = [v_1 \ v_2 \ \dots \ v_n] = [e_{1,1} \ e_{2,1} \ \dots \ e_{S^M,1} \ e_{1,2} \ \dots \ e_{S^M,Q}] \quad (19)$$

o vetor parâmetro é:

$$x^T = [x_1 \ x_2 \ \dots \ x_n] = [w_{1,1}^1 \ w_{1,2}^1 \ \dots \ w_{S^1,R}^1 \ b_1^1 \ \dots \ b_{S^1}^1 \ w_{1,1}^2 \ \dots \ b_{S^2}^M] \quad (20)$$

$$N = Q \times S^M$$

e

$$n = S^1(R + 1) + S^2(S^1 + 1) + \dots + S^M(S^{M-1} + 1)$$

Portanto, fazendo essas substituições em Eq. (7), a matriz Jacobiana para treinamento em redes multicamadas, pode ser escrita:

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial e_{1,1}}{\partial w_{1,1}^1} & \frac{\partial e_{1,1}}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_{1,1}}{\partial w_{S^1,R}^1} & \frac{\partial e_{1,1}}{\partial b_1^1} & \cdots \\ \frac{\partial e_{2,1}}{\partial w_{1,1}^1} & \frac{\partial e_{2,1}}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_{2,1}}{\partial w_{S^1,R}^1} & \frac{\partial e_{2,1}}{\partial b_1^1} & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \\ \frac{\partial e_{S^M,1}}{\partial w_{1,1}^1} & \frac{\partial e_{S^M,1}}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_{S^M,1}}{\partial w_{S^1,R}^1} & \frac{\partial e_{S^M,1}}{\partial b_1^1} & \cdots \\ \frac{\partial e_{1,2}}{\partial w_{1,1}^1} & \frac{\partial e_{1,2}}{\partial w_{1,2}^1} & \cdots & \frac{\partial e_{1,2}}{\partial w_{S^1,R}^1} & \frac{\partial e_{1,2}}{\partial b_1^1} & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \end{bmatrix}.$$

(21)

Os termos nesta matriz jacobiana, podem ser computados por uma simples modificação no algoritmo de *backpropagation*.

O algoritmo *backpropagation* padrão calcula termos como:

$$\frac{\partial F'(x)}{\partial x} = \frac{\partial e_q^T e_q}{\partial x_1} \quad (22)$$

Para os elementos da matriz Jacobiana necessários para o algoritmo Levenberg-Marquardt, precisa-se calcular termos como:

$$[J]_{h,l} = \frac{\partial v_h}{\partial x_1} = \frac{\partial e_{k,q}}{\partial x_1} \quad (23)$$

Utilizam-se as seguintes derivadas que também são usadas no *backpropagation* padrão:

$$\frac{\partial F'}{\partial w_{i,j}^m} = \frac{\partial F'}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial w_{i,j}^m} \quad (24)$$

onde o primeiro termo do lado direito foi definido como a sensibilidade:

$$s_i^m \equiv \frac{\partial F'}{\partial n_i^m} \quad (25)$$

No processo *backpropagation* calculou as sensibilidades através de uma relação de recorrência desde a última camada para trás até a primeira camada. Pode-se usar o mesmo conceito para calcular os termos necessários para a matriz Jacobiana (Eq.(21)) se define-se uma nova sensibilidade Marquardt:

$$s_{i,h}^m \equiv \frac{\partial v_h}{\partial n_{i,q}^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m} \quad (26)$$

onde, a partir de Eq. (19), $h = (q - 1)S^M + k$

Agora podemos computar elementos do Jacobiano por:

$$[J]_{h,l} = \frac{\partial v_h}{\partial x_1} = \frac{\partial e_{k,q}}{\partial w_{i,j}^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m} \times \frac{\partial n_{i,q}^m}{\partial w_{i,j}^m} = s_{i,h}^m \times \frac{\partial n_{i,q}^m}{\partial w_{i,j}^m} = s_{i,h}^m \times a_{j,q}^{m-1} \quad (27)$$

ou se x_1 é um enviesamento:

$$[J]_{h,l} = \frac{\partial v_h}{\partial x_1} = \frac{\partial e_{k,q}}{\partial b_i^m} = \frac{\partial e_{k,q}}{\partial n_{i,q}^m} \times \frac{\partial n_{i,q}^m}{\partial b_i^m} = s_{i,h}^m \times \frac{\partial n_{i,q}^m}{\partial b_i^m} = s_{i,h}^m \quad (28)$$

Para as sensibilidades Marquardt na camada final, se tem que:

$$s_{i,h}^M = \frac{\partial v_h}{\partial n_{i,q}^M} = \frac{\partial e_{k,q}}{\partial n_{i,q}^M} = \frac{\partial (t_{k,q} - a_{k,q}^M)}{\partial n_{i,q}^M} = -\frac{\partial a_{k,q}^M}{\partial n_{i,q}^M} \\ = \begin{cases} -f^M(n_{i,q}^M) & \text{for } i = k \\ 0 & \text{for } i \neq k \end{cases} \quad (29)$$

Portanto, quando a entrada $\mathbf{p}_q$ foi aplicada à rede e a saída de rede correspondente $\mathbf{a}_q^M$ foi computada, o *backpropagation* Levenberg-Marquardt é inicializado com:

$$S_q^M = -\dot{F}^M(n_q^M) \quad (30)$$

Cada coluna da matriz S_q^M deve ser retropropagada através da rede para produzir uma fila da matriz Jacobiana. As colunas também podem ser retropropagadas em conjunto, utilizando:

$$S_q^m = \dot{F}^m(n_q^m)(W^{m+1})^T S_q^{m+1} \quad (31)$$

As matrizes totais de sensibilidade Marquardt para cada camada são então criadas através do aumento das matrizes calculadas para cada entrada:

$$S^m = [S_1^M | S_2^M | \dots | S_Q^m] \quad (32)$$

Note que para cada entrada que é apresentada à rede, retropropagam-se os vetores S de sensibilidade. Isto é porque calcula-se as derivadas de cada erro individual, em vez da derivada da soma dos quadrados dos erros. Para cada entrada aplicada à rede, haverá S erros (um para cada elemento da saída da rede). Para cada erro, haverá uma linha da matriz Jacobiana.

Após as sensibilidades terem sido retropropagadas, a matriz jacobiana é computada usando Eq. (27) e Eq. (28).

As iterações do algoritmo Levenberg-Marquardt *Backpropagation* (LMBP) podem ser resumidas como se segue:

1. Apresentar todas as entradas na rede e calcular as saídas de rede correspondentes e os erros $e_q = t_q - a_q^M$. Calcular a soma dos erros quadrados sobre todas as entradas $F(\mathbf{x})$, usando Eq. (18).
2. Calcule a matriz Jacobiana, Eq. (21). Calcular as sensibilidades com as relações de recorrência Eq. (31), após inicializar com Eq. (30). Aumentar as matrizes individuais para as sensibilidades Marquardt usando Eq. (32). Calcular os elementos da matriz Jacobiana com Eq. (27) e Eq. (28).
3. Resolver a Eq. (16) para obter Δx_k .
4. Recalcular a soma dos erros quadrados usando $x_k + \Delta x_k$. Se esta nova soma de quadrados for menor do que a calculada no passo 1,

então divide-se μ por ϑ , deixe-se $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k$ e volte ao passo 1. Se a soma dos quadrados não for reduzida, então multiplica-se μ por ϑ e retorna-se ao passo 3

Presume-se que o algoritmo tenha convergido quando a norma do gradiente, Eq. (6), é inferior a algum valor predeterminado, ou quando a soma dos quadrados foi reduzida até ser inferior ao valor de erro objetivo prefixado (HAGAN, 2014).

3.2 CASCADE-FORWARD NEURAL NETWORK(CFNN)

A rede neural *Cascade-Forward* é uma classe de rede neural que é semelhante às redes *Feedforward*, mas cada camada inclui uma conexão a partir da entrada e uma conexão para cada uma das camadas antes dela, pode-se observar estas conexões na Figura 2.

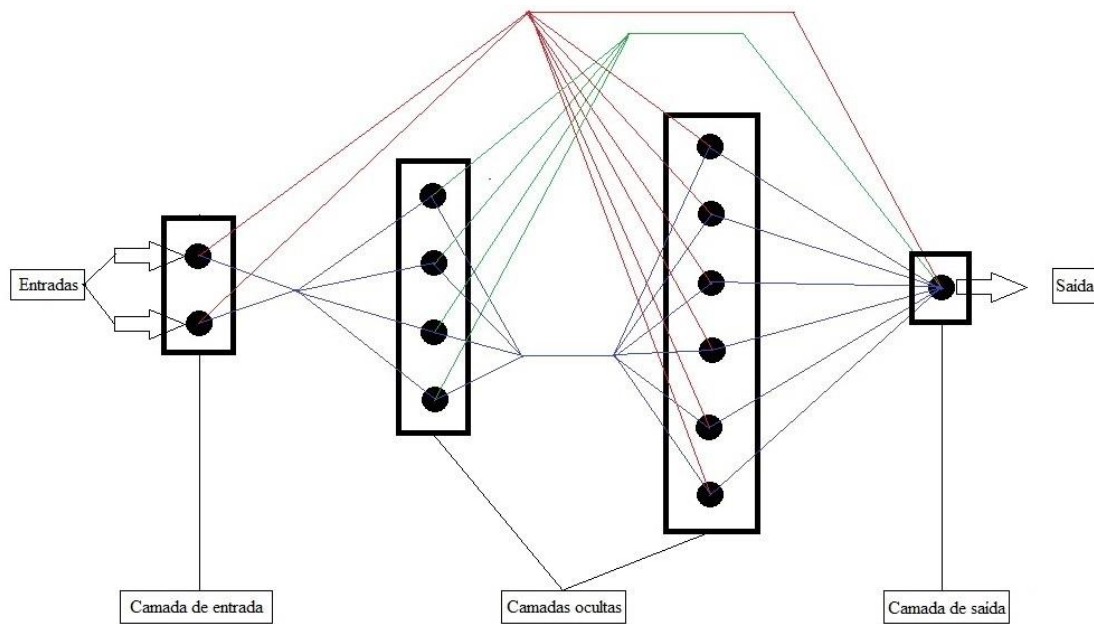


Figura 2 – Arquitetura do *Cascade-Forward*

Fonte: Elaboração própria.

Como nas redes de *Feedforward*, uma rede *Cascade-Forward* de duas ou mais camadas pode aprender qualquer relação finita de entrada-saída. A rede neural *Cascade-Forward* pode ser usada para qualquer tipo de mapeamento de entrada e de saída.

A vantagem deste método é que ele acomoda a relação não linear entre entrada e saída, não eliminando a relação linear entre os dois.

As equações são formadas do modelo *CFNN* pode ser escrito como se segue:

$$y = \sum_{i=1}^n f^i w_i^i x_i + f^o \left(\sum_{j=1}^k w_j^o f_j^h \left(\sum_{i=1}^n w_{ji}^h x_i \right) \right) \quad (33)$$

Sendo: f^i é a função de ativação da camada de entrada para a camada de saída, w_i^l é o peso da camada de entrada para a camada de saída, f^o é a função de ativação na camada de saída e f_j^h é uma função de ativação na camada oculta. Se um bias w_b é adicionado à camada de entrada e a função de ativação de cada neurônio na camada oculta é f^h , então, a equação se torna

$$y = \sum_{i=1}^n f^i w_i^l x_i + f^o (w^b + \sum_{j=1}^k w_j^o f_j^h (w^b + \sum_{i=1}^n w_{ji}^h x_i)) \quad (34)$$

A Figura 2 mostra que existe uma relação direta entre entrada e saída. A consequência desta forma de relacionamento é que o peso da rede a ser estimado aumenta tanto quanto os neurônios na camada de entrada. Como no *FFNN*, o algoritmo de treinamento é Levenberg-Marquardt *Backpropagation* (*LMBP*). (WARSITO; SANTOSO, 2018)

3.3 GENERALIZED REGRESSION NEURAL NETWORK(*GRNN*)

GRNN, proposta por Specht (1991), é uma rede capaz de estimar uma variável dependente de uma variável independente através de um conjunto finito de variáveis. Seu tipo de treinamento é supervisionado, tem uma estrutura não recorrente e é realizada numa única apresentação dos padrões de treinamento. Esta rede tem uma estrutura fixa, dada por uma camada de entrada, duas camadas intermédias e uma camada de saída, onde o número de neurônios em cada camada depende do número de padrões de treinamento, entradas e saídas da rede. A sua capacidade de generalização é mantida mesmo com a utilização de dados escassos, e uma vez que se baseia numa função de densidade de probabilidades, as probabilidades de queda para mínimos locais são muito baixas (NIU, 2005).

Esta rede tem sido amplamente utilizada nos últimos anos, principalmente para problemas relacionados com a previsão de séries temporais, tais como cargas elétricas de curto prazo (NIU, 2005) (NOSE-FILHO, 2009) (TRIPATHI, 2008) (XIA, 2010) (YAP, 2006). Foram propostas várias modificações para melhorar a sua capacidade de generalização. No artigo de Niu no 2005, os autores propõem a pesquisa do parâmetro de dispersão usando um algoritmo evolutivo diferencial modificado. O trabalho de Yap no

2006 combina a utilização de redes baseadas na teoria da ressonância adaptativa (ART) (GROSSBERG, 2013). com *GRNN*.

A *GRNN* convencional foi programada com o auxílio das toolboxes das RNAs do MATLAB por meio da função *newgrnn*. A regressão consiste em estimar uma variável dependente a partir de uma variável independente por meio de um conjunto finito de variáveis e que pode ser visto na figura 3. Na teoria dos sistemas, a variável dependente corresponde à saída e a variável independente corresponde à entrada e pode ser descrita por meio de vectores, cujo número de elementos indica o número de entradas e saídas do sistema.

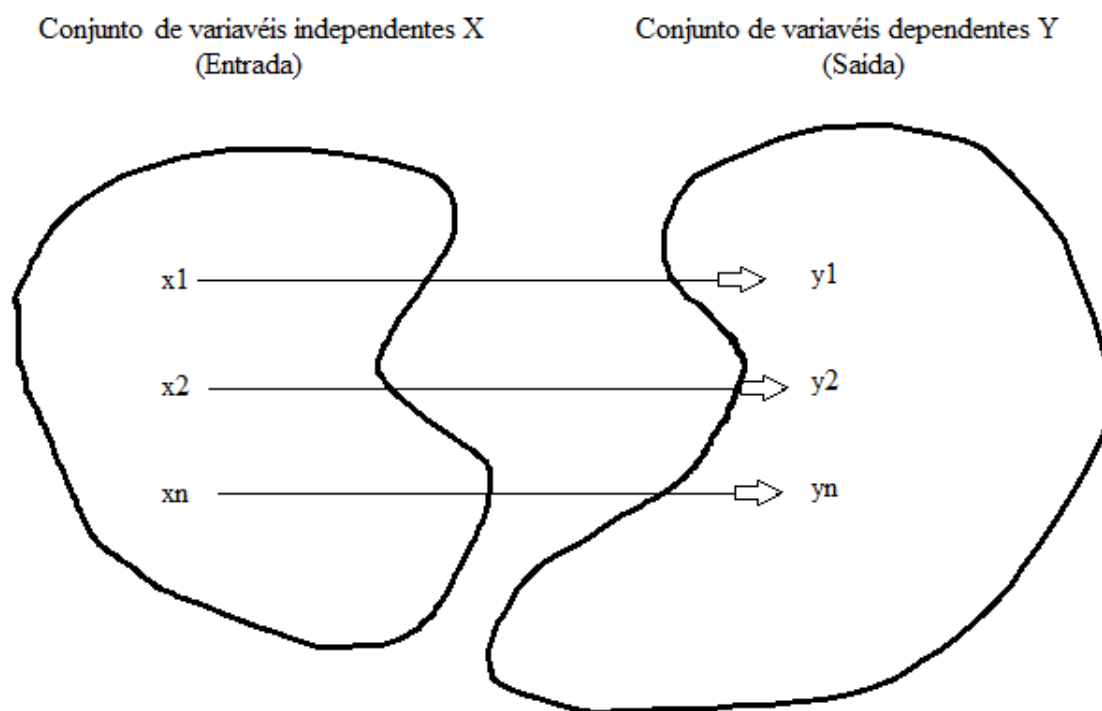


Figura 3 – Conjuntos de variáveis independentes X é conjuntos de variáveis independentes Y

Fonte: Elaboração própria.

Considerando dois conjuntos finitos, X e Y de variáveis independentes (entradas) e dependentes (saídas) e sendo

uma variável de entrada desconhecida, a saída pode ser estimada pela regressão por meio das variáveis contidas nos conjuntos $\mathbf{X}$ e $\mathbf{Y}$, conforme a:

$$E(Y|\hat{x}) = \frac{\int_{-\infty}^{+\infty} yf(\hat{x},Y)dy}{\int_{-\infty}^{+\infty} f(\hat{x},Y)dy} \quad (35)$$

A função densidade de probabilidade $f(\hat{x}, Y)$ pode ser obtida por meio dos conjuntos $\mathbf{X}$ e $\mathbf{Y}$. A rede neural de regressão generalizada utiliza como função densidade de probabilidade a estimação de Kernel, também conhecida como janela de Parzen. A estimativa da variável $\hat{y}$ dado uma variável desconhecida $\hat{x}$ pode ser determinada por meio da Eq. (35), sendo n o número de amostras dos conjuntos $\mathbf{X}$ e $\mathbf{Y}$, y_i a variável dependente associada à independente x_i e a_i o valor da i -ésima distância euclidiana da variável desconhecida $\hat{x}$ com as variáveis x_i , dadas por (Eq. (36)).

$$\hat{y}(\hat{x}) = \frac{\sum_{i=1}^n y_i \exp\left(\frac{-a_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(\frac{-a_i^2}{2\sigma^2}\right)} \quad (36)$$

$$a_i = (\hat{x} - x_i)^T (\hat{x} - x_i) \quad (37)$$

A função de transferência da *GRNN* é uma função gaussiana dada por Eq. (37). O parâmetro σ , denominado *spread*, define a abertura desta função. A função de transferência utilizada pelo *toolbox* de RNAs do MATLAB é uma função do tipo gaussiana (*radbas*) dada por Eq. (38), ilustrada pela Figura 4.

$$f(a_i) = \exp\left(\frac{-a_i^2}{2\sigma^2}\right) \quad (38)$$

$$f(a_i) = \exp\left(\frac{-0.8326a_i^2}{\sigma^2}\right) \quad (39)$$

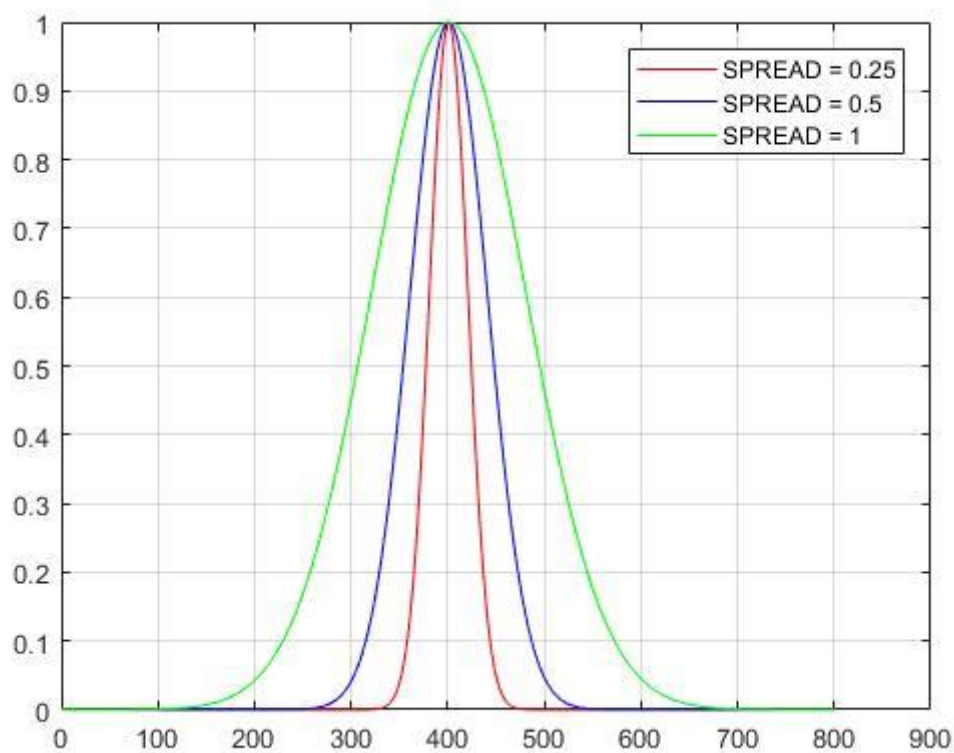


Figura 4 – Valor do parâmetro de dispersão (*Spread*)

Fonte :Elaboração própria.

O valor do parâmetro de dispersão ou *spread* define a área da vizinhança e, consequentemente, o número de padrões a ter em conta para a estimativa de uma variável. Como pode-se observar na Figura 4 quanto maior for o valor do *spread*, maior será a área da vizinhança, e mais padrões serão tidos em conta e quanto mais baixo for este valor, menor será a área da vizinhança, e menos padrões serão tidos em conta. O valor do *spread* está compreendido entre 0 e 1.

A *GRNN* possui uma estrutura fixa, disposta em quatro camadas, uma de entrada, duas intermediárias e uma de saída, conforme Figura 5. O número de neurônios de cada camada depende exclusivamente do número das entradas, saídas e de amostras do conjunto de dados de treinamento.

O padrão de entrada da rede é representado por um vetor coluna $\hat{X}$ *ns*-dimensional e o padrão de saída é representado por um vetor coluna $\hat{Y}$, *ns*-dimensional.

O conjunto de dados de treinamento é representado por duas matrizes, uma matriz $X_{ne \times n}$ contendo os padrões de entrada e outra matriz $Y_{ns \times n}$ contendo os padrões de saída. O número n de colunas de cada matriz representa a quantidade de amostras do conjunto de dados de treinamento.

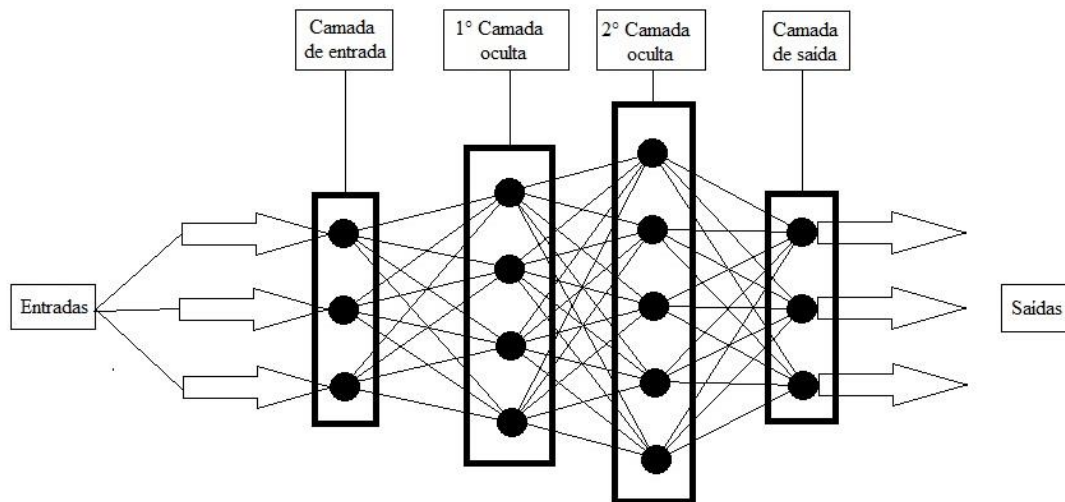


Figura 5 – Arquitetura do *GRNN*

Fonte: Elaboração própria

A camada de entrada da *GRNN* é composta por ne neurônios. A primeira intermediária é composta por n neurônios. A segunda intermediária é composta por ns neurônios do tipo numerador e um neurônio do tipo denominador. A de saída é composta por ns neurônios de saída.

O algoritmo de treinamento e o algoritmo de diagnóstico/previsão da *GRNN* podem ser dados pelas seguintes etapas:

Treinamento:

1. Leitura das matrizes X e Y e do parâmetro *spread*.

Diagnóstico/Previsão:

1. Leitura do padrão desconhecido $\hat{X}$;
2. Cálculo das distâncias euclidianas, a_i dadas por Eq. (40);
3. Cálculo dos níveis de pertinência, b_i dados por Eq. (41);

4. Cálculo das saídas dos neurônios do tipo numerador, $\mathbf{c}_j$ dadas por Eq. (42), e da saída do neurônio do tipo denominador, $\mathbf{d}$ dada por Eq. (43);
5. Cálculo das saídas estimadas, $\hat{\mathbf{y}}_j$ dadas por Eq. (44);
6. A saída $\hat{\mathbf{Y}}$ é dada por Eq. (45).

$$a_i = (\hat{X} - X_i)^T (\hat{X} - X_i), i = 1, \dots, n \quad (40)$$

$$b_i = \exp\left(\frac{-0.8326a_i^2}{\sigma^2}\right), i = 1, \dots, n \quad (41)$$

$$c_j = \sum_{i=1}^n y_{j,i} b_i, j = 1, \dots, ns \quad (42)$$

$$d = \sum_{i=1}^n b_i \quad (43)$$

$$\hat{y}_j = \frac{c_j}{d}, j = 1, \dots, ns \quad (44)$$

$$\hat{\mathbf{Y}} = [\hat{y}_1 \dots \hat{y}_j \dots \hat{y}_{ns}]^T \quad (45)$$

(NOSE FILHO, 2011).

4 METODOLOGIAS

No início o plano era utilizar fatores macroeconômicos para a previsão do IBEX 35, mas não se conseguiu obter dados diários, o que resultou em uma quantidade insuficiente de dados para esta pesquisa. Então, decidiu-se abortar esta ideia e começar a usar outros dados, como o valor de fechamento de outros índices de ações estrangeiras, commodities e a taxa de câmbio entre duas moedas. O preço de fechamento de um ativo é o último preço em que ele foi negociado numa sessão da bolsa.

Para calcular o valor do dia seguinte, foi utilizada a conhecida técnica do janelamento (GOUK, 2014), que consiste em inserir como entrada os valores dos dias anteriores ao valor que se precisa determinar. Neste caso, se entrariam com os valores de fechamento dos 5 dias anteriores dos índices de ações estrangeiras, taxas de câmbio e matérias-primas e do mesmo índice IBEX 35.

Para inferir a qualidade da previsão, usa-se o conceito *MAPE*, o qual é consolidado na literatura especializada (PARK *et al.*, 1991):

$$MAPE = \frac{1}{NA} \sum_{h=1}^{NA} \left| \frac{y_{atual}(h) - y_{fprevisto}(h)}{y_{atual}(h)} \right| \times 100 \quad (46)$$

sendo:

MAPE : *Mean Absolute Percent Error* (percentual do erro absoluto médio).

NA : número da amostragem da série de tempo discreto;

$y_{atual}(h)$: valor de referência (para fins de comparação) relativa ao instante discreto h ;

$y_{previsto}(h)$: valor previsto referente ao instante discreto h .

Considerando que um valor *MAPE* aceitável deve ser inferior a 20% e um excelente valor inferior a 10%. (GILLILAND, 2010).

4.1 OBTENÇÃO DAS ENTRADAS

Os diferentes valores de fechamento dos diferentes índices foram obtidos na seção de finanças do *Yahoo*, que é junto com o *Investing* de uma das páginas mais conceituadas para encontrar dados históricos sobre índices de ações. A partir de ambas as páginas é possível obter dados diários que podem ser facilmente importados em formato Excel.

Ao obter os dados históricos dos diferentes índices bolsistas para usar como entrada, devem-se usar todos eles da mesma página, visto que a princípio estava-se pensando em usar tanto o *Yahoo Finance* quanto o *Investing*, isto porque nestas entidades estavam faltando dados de alguns índices bolsistas. Mas não é recomendável fazer isso porque os valores do *Yahoo* e do *Investing* não podem ser misturados, pois seus valores de fechamento não são os mesmos e os valores de correlação linear entre os índices dão valores errados. Então, deve-se escolher os parâmetros só de uma página. Neste caso particular, foi decidido utilizar o *Yahoo Finance*.

A princípio, os dados históricos foram tomados de janeiro de 2010 a dezembro de 2019 dos mais importantes índices bolsistas do mundo e se comparou sua correlação linear com o índice IBEX 35, descartando aqueles com correlação que considerava baixa, tomando como válidos apenas aqueles cujo valor absoluto de sua correlação linear com o índice IBEX 35 fossem o maior possível.

Portanto, em uma primeira seleção, os índices bolsistas das bolsas estrangeiras escolhidos foram os seguintes 12: AEX (Holanda), BEL 20 (Bélgica), CAC 40 (França), DAX 30 (Alemanha), HANG SENG (Hong Kong), NIKKEI 225 (Japão), OMXS 30 (Suécia), S&P ASX 200 (Austrália) S&P TSX 60 (Canadá), SMI 30 (Suíça), Taiwan Weighted (Taiwan) e o próprio IBEX 35 (Espanha).

Os três seguintes foram selecionados como matérias primas influenciadoras: Milho, Ouro e Prata.

Finalmente, os quatro seguintes índices foram escolhidos como taxas de câmbio: USD/JPY (taxa de câmbio entre o dólar americano e o iene japonês), EUR/CAD (taxa de câmbio entre o euro e o dólar canadense), EUR/AUD (taxa de câmbio entre o

euro e o dólar australiano) e EUR/JPY (taxa de câmbio entre o euro e o iene japonês). Dando um total de 19 índices.

Com o propósito de analisar a robustez das redes neurais, verifica-se se as RNAs treinadas com os dados de 2010 a 2019 seriam capazes de prever com precisão os valores de fechamento do ano 2020, que como se sabe foi um ano em que ocorreu a pandemia global e, portanto, o mercado de ações sofreu quedas significativas.

4.2 PROCESAMENTO DAS ENTRADAS

Para realizar o processo de previsão do índice IBEX 35, para o dia seguinte, foi usada a técnica de janelamento que já foi descrita anteriormente, que consiste em colocar os dados dos dias anteriores em relação aos dados que se quer conhecer.

Neste caso, foi obtida a previsão usando as redes neurais descritas anteriormente, diferenciando 2 casos:

1. No primeiro caso, foram utilizados como entrada apenas os dados de 5 dias anteriores ao valor do dia do mesmo índice IBEX 35 que quer se conhecer, dando 5 entradas para o cálculo.
2. No segundo caso, foram usados, como dados de entrada, 19 índices que possuem a maior correlação possível com o IBEX 35, considerando-se somente aqueles que possuem correlação. Neste caso, apartam-se somente 6 índices (entre os 19 indicados) que somados ao mesmo IBEX 35 daria um total de 7 índices. Portanto, levando em conta os valores dos 5 dias anteriores, daria um total de 35 entradas. Os 6 índices escolhidos seriam os índices bolsistas Hang Seng (Hong Kong) é BEL 20(Bélgica), as matérias-primas Maiz é Oro e as taxas de câmbio EUR/JPY é EUR/AUD.

A ideia de realizar 2 casos diferentes é ver a diferença de desempenho das previsões quando são colocadas diferentes entradas, nas diferentes redes neurais aplicando três RNAs diferentes.

Finalmente, como mencionado anteriormente, é analisada a robustez das redes neurais já treinadas, verificando como a rede é capaz de prever os valores de fechamento para 2020.

O procedimento consiste simplesmente em simular os valores nas RNA já treinadas dos cinco dias anteriores (janelamento) do ano 2020, do valor do dia de fechamento a ser conhecido.

A robustez será analisada para ambos os casos (caso 1 e caso 2) sendo 5 entradas para o caso 1 e 35 entradas para o caso 2.

5 RESULTADOS OBTIDOS.

Na sequência, são mostradas tabelas de resultados do *MAPE*. A métrica *MAPE* é usada visando quantificar a precisão de uma técnica de previsão comparando-se os valores desejados com os valores calculados. São empregadas, na análise, as três RNAs mencionadas anteriormente: *Feedforward Neural Network (FNN)*, *Cascade-Forward Neural Network (CFNN)*, e *Generalized Regression Neural Network (GRNN)*.

Quando são utilizadas o *Feedforward* e o *Cascade-Forward*, são mostrados os diferentes resultados obtidos quando se varia o número de neurônios da camada oculta e sua disposição e, também, quando se muda a porcentagem de treinamento.

Tanto no caso de *Feedforward* como no *Cascade-Forward* foram usados o algoritmo de treinamento *Levenberg-Marquardt Backpropagation* (MARQUARDT, 1963). A função de transferência usada foi a sigmoide tangente hiperbólica.

Quanto às porcentagens de treinamento, foram utilizadas três variantes:

1. 70% de treinamento.
2. 80% de treinamento.
3. 90% de treinamento.

Em relação à distribuição da camada oculta, foram usados dois tipos de distribuição:

1. 1 camada oculta.
2. 3 camadas ocultas.

Quando é utilizada a *GRNN*, também, foram utilizadas três variantes, mas, neste caso, como já foi mencionado anteriormente, tem-se de fazer, manualmente, usando-se uma porcentagem da mostra de dados para treinar.

1. 70% de treinamento.
2. 80% de treinamento.
3. 90% de treinamento.

Quando se usa a *GRNN*, o único parâmetro que foi mudado é a dispersão (*spread*), isto porque o número de neurônios ocultos é igual ao número de amostras do treinamento.

1. *Spread* = 0.1
2. *Spread* = 1.

Nas tabelas a seguir, a célula com o menor valor do *MAPE* em cada linha é registrada na cor vermelha e amostrada em porcentagem. Após cada tabela de resultados, são mostrados os gráficos que comparam, visualmente, a saída desejada (*Target*) em vermelho e a saída calculada pela rede neural (*Output*) em azul, correspondendo, em cada caso, à configuração com valor de *MAPE* menor.

Cada rede neural treinada (com os parâmetros anteriormente referidos), poderia prever os valores de fechamento do IBEX 35 desde 01/01/2020 até o 30/09/2020. Trata-se de análise da robustez das RNAs treinadas na predição dos valores de ano 2020, em um ano que por causa da pandemia as bolsas de valores internacionais tiveram muitas flutuações por consequência da crise global.

5.1 ANÁLISE DO MAPE DO CASO 1

Neste caso, como informado anteriormente, foram usados apenas os dados dos 5 dias anteriores do IBEX 35 como entrada.

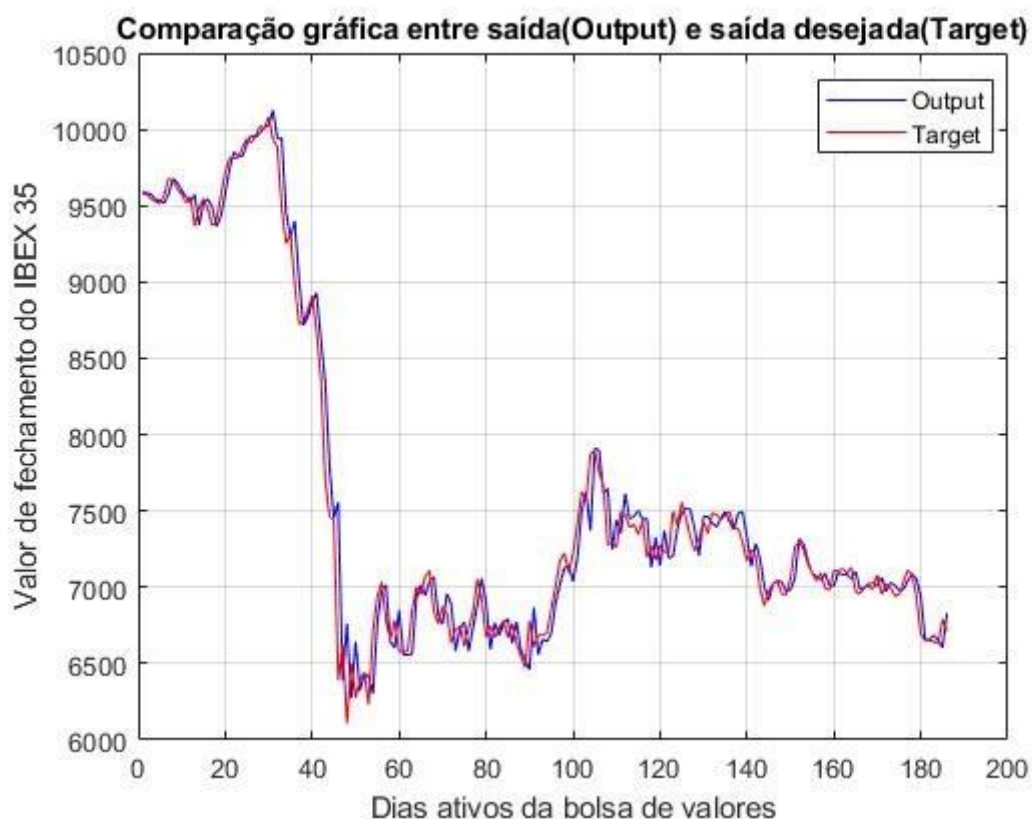
5.1.1 USANDO A REDE NEURAL *FEEDFORWARD*

Tabela 1 - Resultados do *MAPE* da rede neural *Feedforward* para o primeiro caso com a configuração 70% treinamento, relativo à predição dos valores de fechamento do ano 2020.

FEEDFORWARD														
INPUT: IBEX 35														
70%														
1 camada oculta							3 camadas ocultas							
Neurônios por camada							Neurônios por camada							
100	150	200	250	300	400	450	7-5-3	2-3-4	8-6-4	16-12-8	48-36-24	6-5-4	4-5-6	3-5-7
1.934	2.189	2.509	1.946	3.223	3.012	3.046	1.667	1.619	1.667	1.855	1.818	13.674	1.669	1.612

Fonte: Elaboração própria

Figura 6 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (3,5,7) neurônios, relativo à predição dos valores de fechamento do ano 2020.



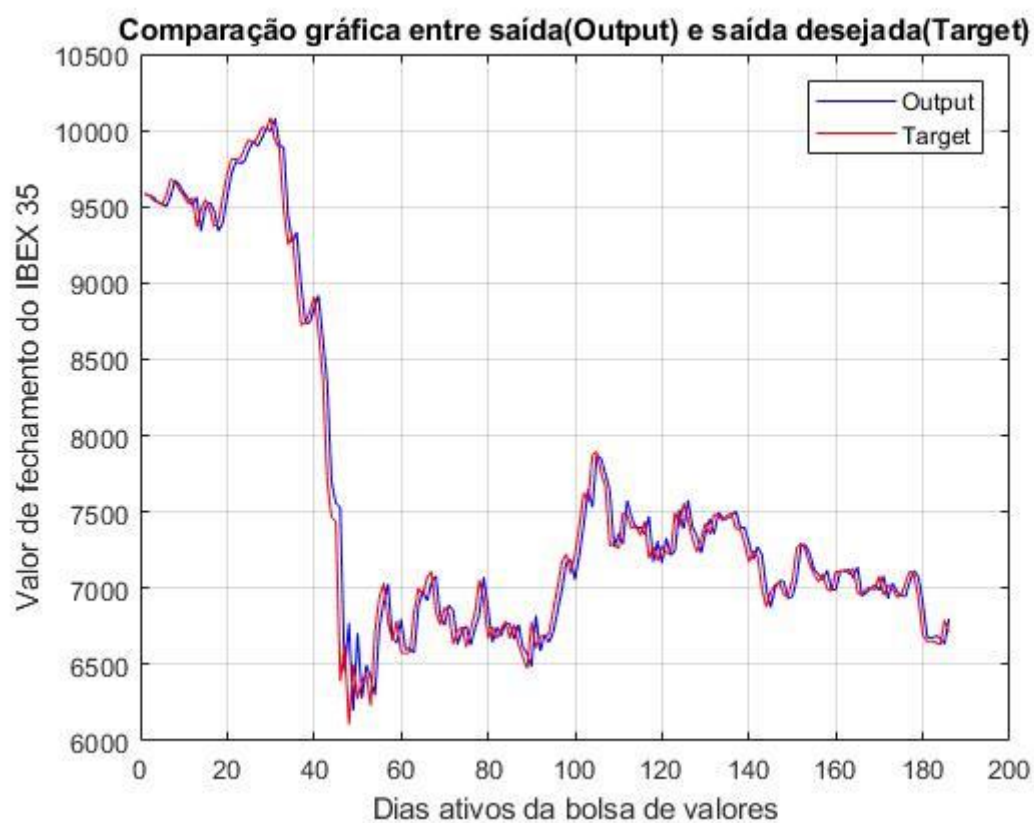
Fonte: Elaboração própria

Tabela 2 - Resultados do *MAPE* da *rede neural Feedforward* para o primeiro caso com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

FEEDFORWARD															
INPUT:IBEX 35															
80%															
1 camada oculta							3 camadas ocultas								
Neurónios por camada							Neurónios por camada								
100	150	200	250	300	400	450	7-5-3	2-3-4	8-6-4	16-12-8	48-36-24	6-5-4	4-5-6	3-5-7	4-3-2
1.898	2.548	2.561	2.463	2.623	4.254	4.044	1.635	1.593	1.633	1.665	1.834	1.988	1.600	1.689	1.605

Fonte: Elaboração própria

Figura 7 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (2,3,4) neurônios, relativo à predição dos valores de fechamento do ano 2020.



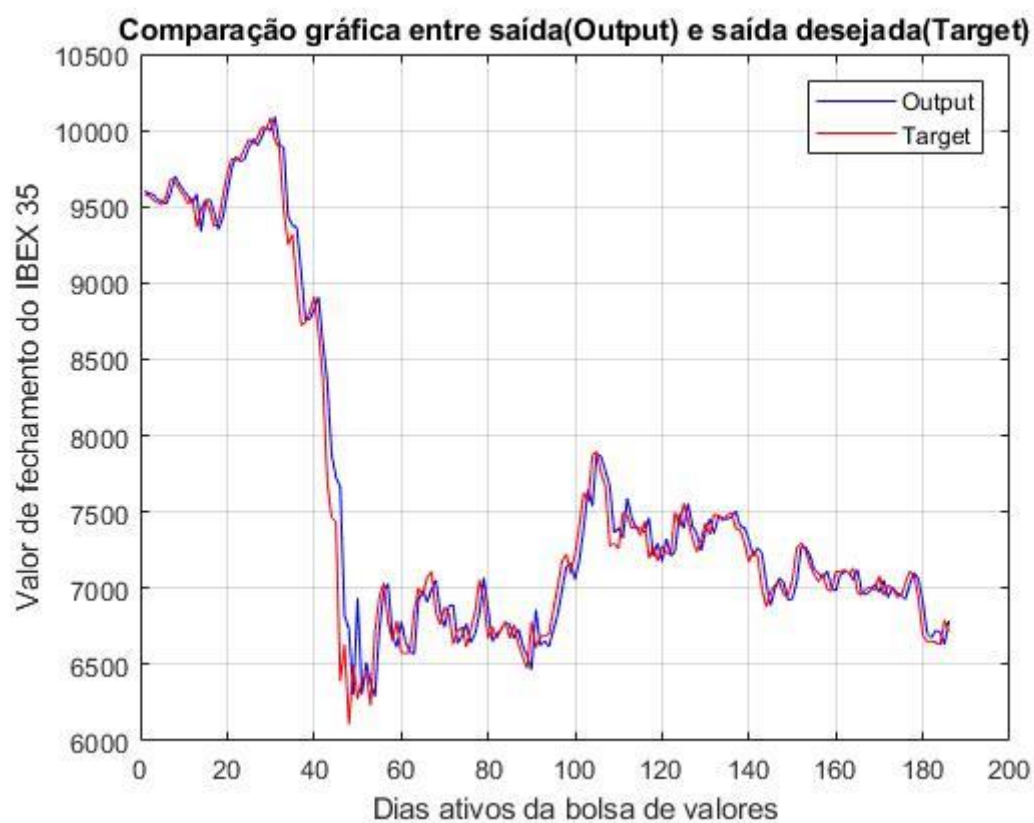
Fonte: Elaboração própria

Tabela 5 - Resultados do *MAPE* da rede neural *Cascade-Forward* para o primeiro caso com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

CASCADE														
INPUT: IBEX 35														
80%														
1 camada oculta							3 camadas ocultas							
Neurônios por camada							Neurônios por camada							
100	150	200	250	300	400	450	7-5-3	2-3-4	8-6-4	16-12-8	48-36-24	6-5-4	4-5-6	3-5-7
2.129	2.766	2.199	2.584	3.410	3.941	3.045	1.743	1.973	1.835	1.657	2.034	1.873	1.776	1.667
														2.023

Fonte: Elaboração própria

Figura 10 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (16,12,8) neurônios, relativo à predição dos valores de fechamento do ano 2020.



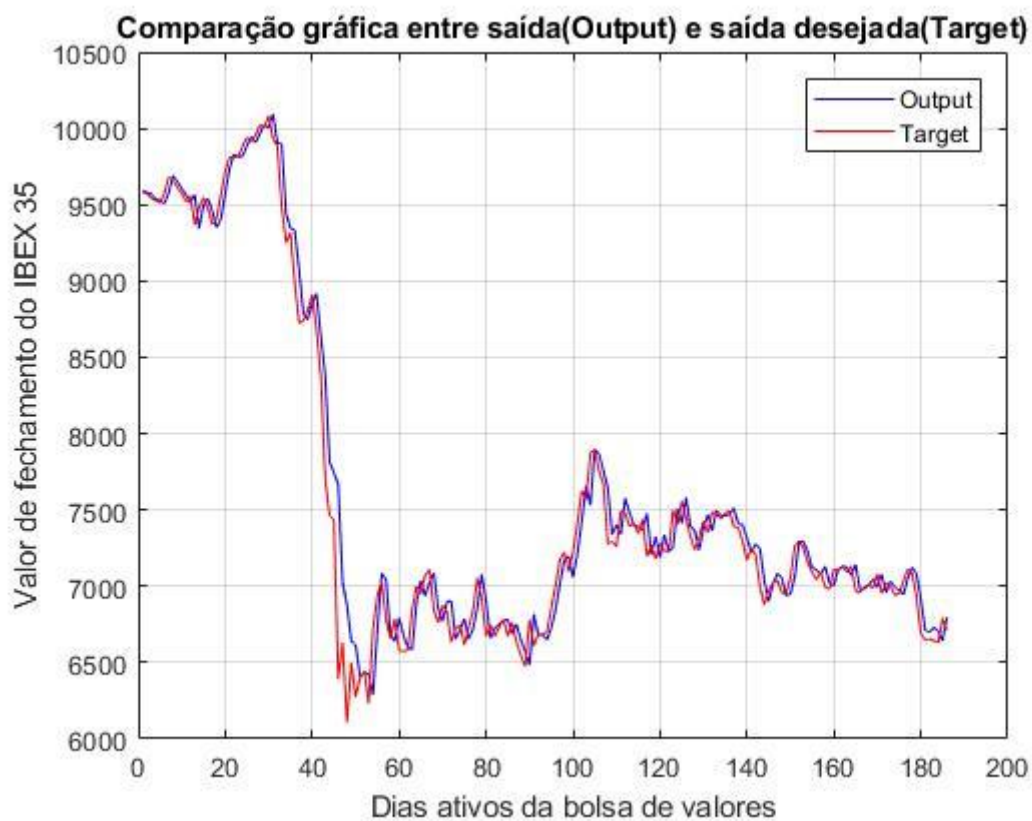
Fonte: Elaboração própria

Tabela 6 - Resultados do *MAPE* da rede neural *Cascade-Forward* para o primeiro caso com a configuração 90% treinamento, relativo à predição dos valores de fechamento do ano 2020.

CASCADE														
INPUT: IBEX 35														
90%														
1 camada oculta							3 camadas ocultas							
Neurônios por camada							Neurônios por camada							
100	150	200	250	300	400	450	7-5-3	2-3-4	8-6-4	16-12-8	48-36-24	6-5-4	4-5-6	3-5-7
2.134	1.930	3.611	2.480	3.510	2.961	3.239	1.665	1.926	2.015	1.654	1.974	1.813	1.660	1.647

Fonte: Elaboração própria

Figura 11 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (3,5,7) neurônios, relativo à predição dos valores de fechamento do ano 2020.



Fonte: Elaboração própria

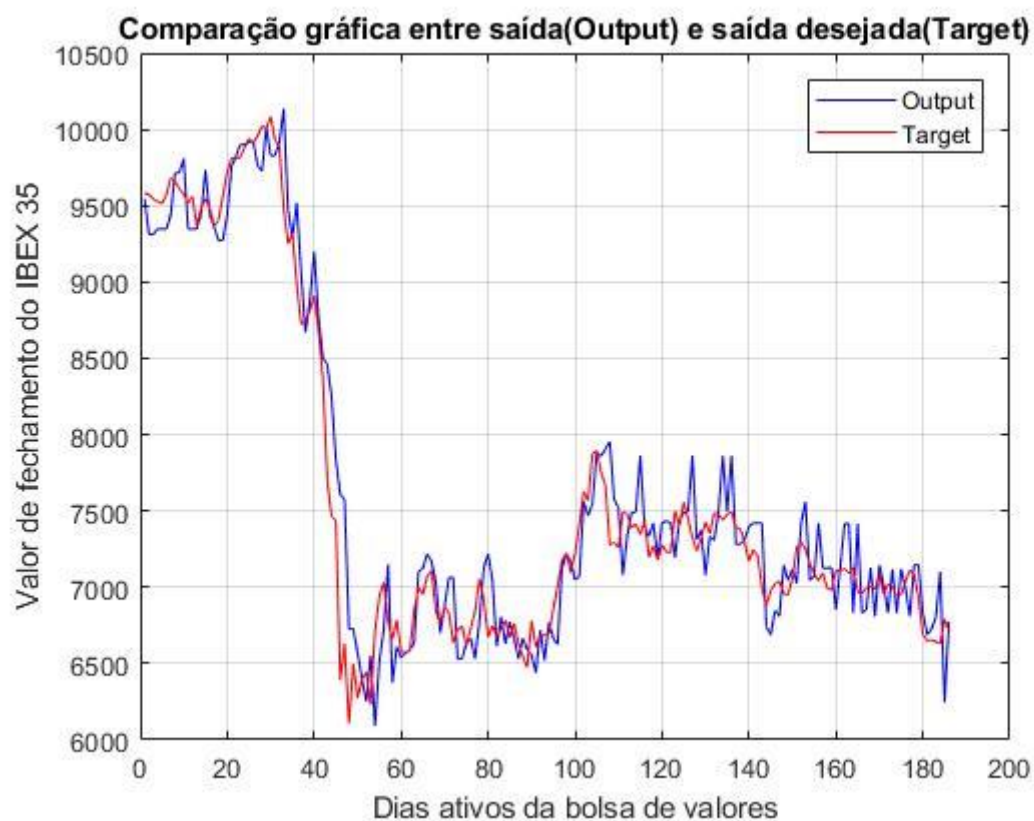
5.1.3 USANDO A *GENERALIZED REGRESION NEURAL NETWORK*

Tabela 7 - Resultados do *MAPE* da rede neural *GRNN* para o primeiro caso com a configuração 70% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT:IBEX 35	
70%	
SPREAD	
0.1	1
2.696	2.693

Fonte: Elaboração própria

Figura 12 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



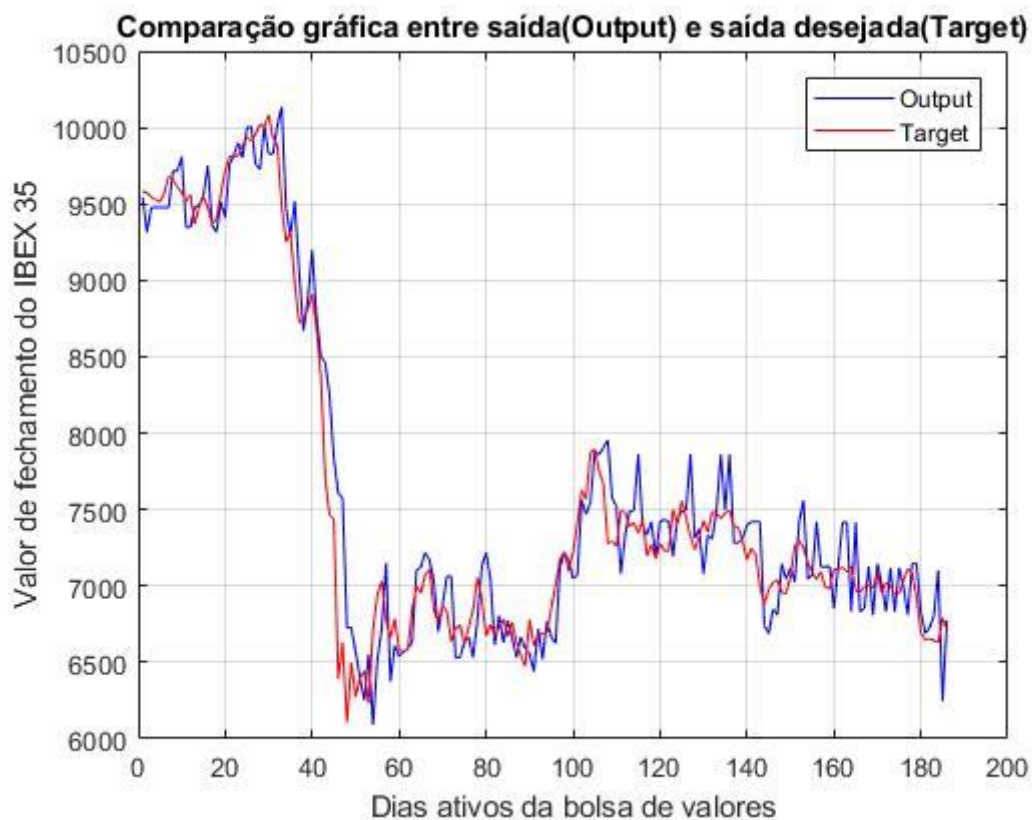
Fonte: Elaboração própria

Tabela 8 - Resultados do *MAPE* da rede neural *GRNN* para o primeiro caso com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT:IBEX 35	
80%	
SPREAD	
0.1	1
2.670	2.668

Fonte: Elaboração própria

Figura 13 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



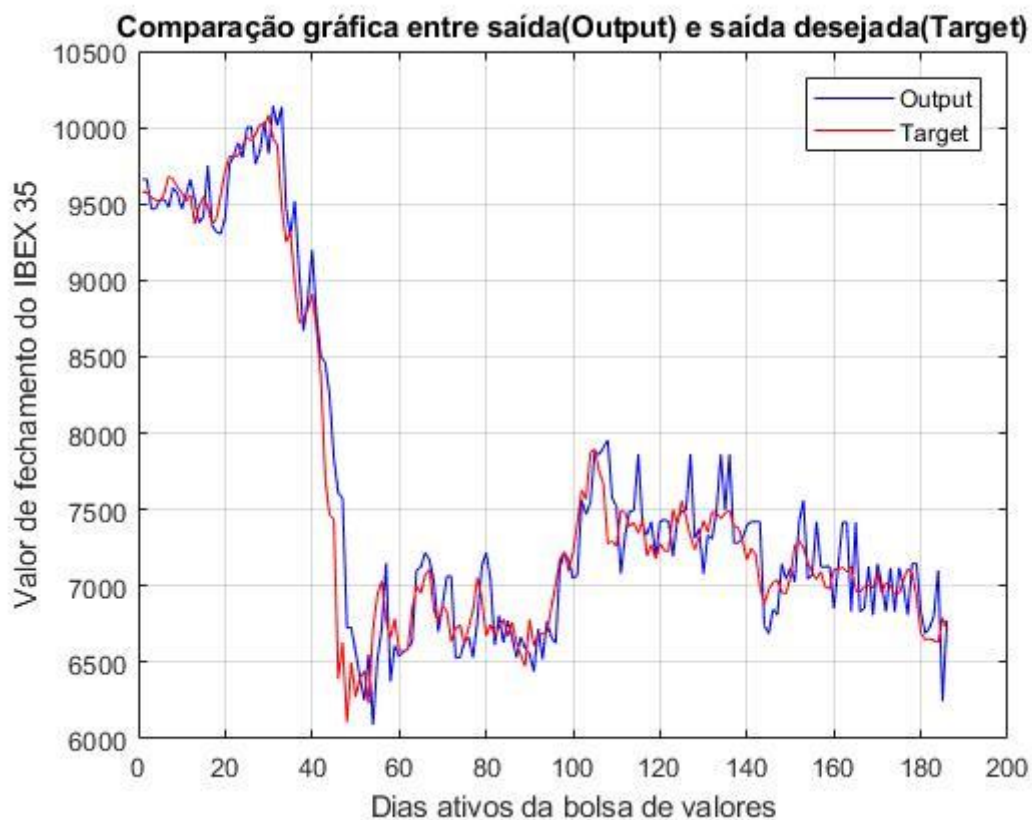
Fonte: Elaboração própria

Tabela 9 - Resultados do *MAPE* da rede neural *GRNN* para o primeiro caso com a configuração 90% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT: IBEX 35	
90%	
SPREAD	
0.1	1
2.663	2.660

Fonte: Elaboração própria

Figura 14 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



Fonte: Elaboração própria

5.2 ANÁLISE DO *MAPE* DO CASO 2

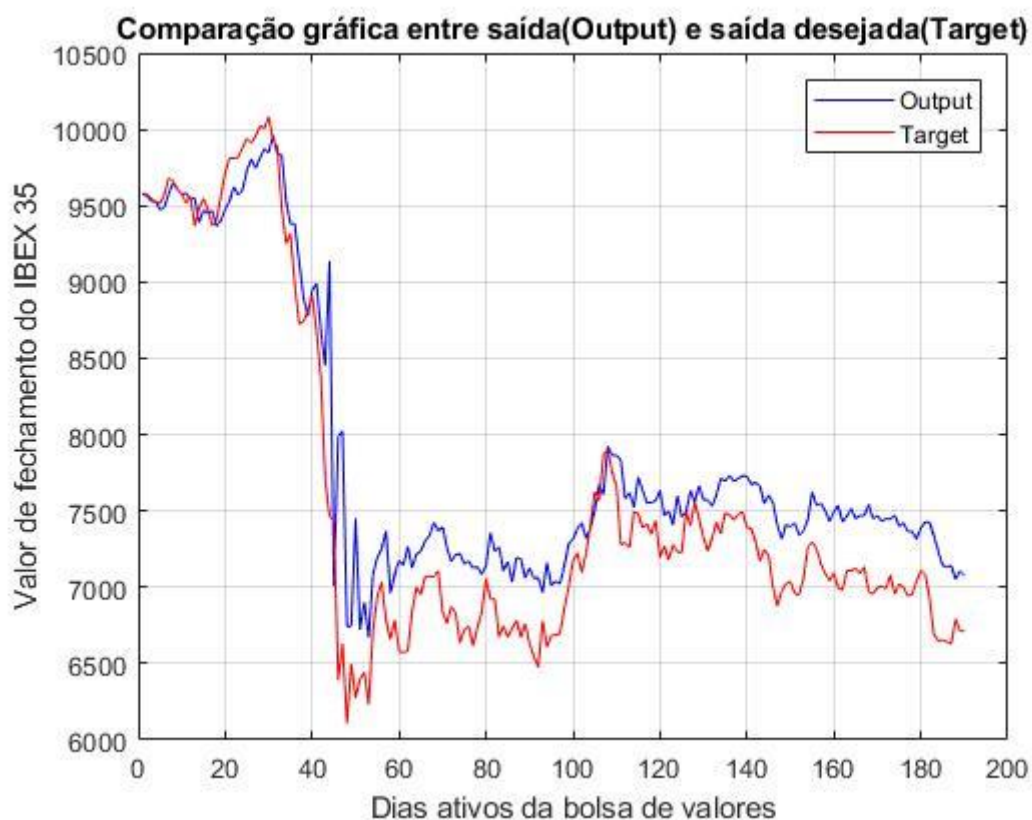
Neste caso, como já se citou anteriormente, tomam-se como entradas os índices que possuem uma correlação com o IBEX 35 compreendida entre 0,3 e 0,8. Devido a que uma correlação muito alta entre os parâmetros de entrada pode produzir erros por causa da redundância. Escolhendo 6 índices que somados ao mesmo IBEX 35 chega-se num total de 7 índices. Portanto, levando em conta os valores dos 5 dias anteriores, totalizam-se 35 entradas.

Tabela 11 - Resultados do *MAPE* da rede neural *Feedforward* (para o segundo caso) com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

FEEDFORWARD																
INPUT:SELEÇÃO DE ÍNDICES																
80%																
1 camada oculta						3 camadas ocultas										
NEURÔNIOS POR CAMADA						NEURÔNIOS POR CAMADA										
100	150	200	250	300	400	48-36-24	50-55-60	40-35-30	20-15-5	70-40-10	25-15-5	30-25-20	30-20-30	10-40-70	30-35-40	20-25-30
19.944	30.250	98.499	43.524	113.687	95.983	17.376	27.171	20.833	9.287	19.198	8.257	15.235	23.769	54.460	16.824	16.186
																4.603

Fonte: Elaboração própria

Figura 16 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (5,15,25) neurônios, relativo à predição dos valores de fechamento do ano 2020.



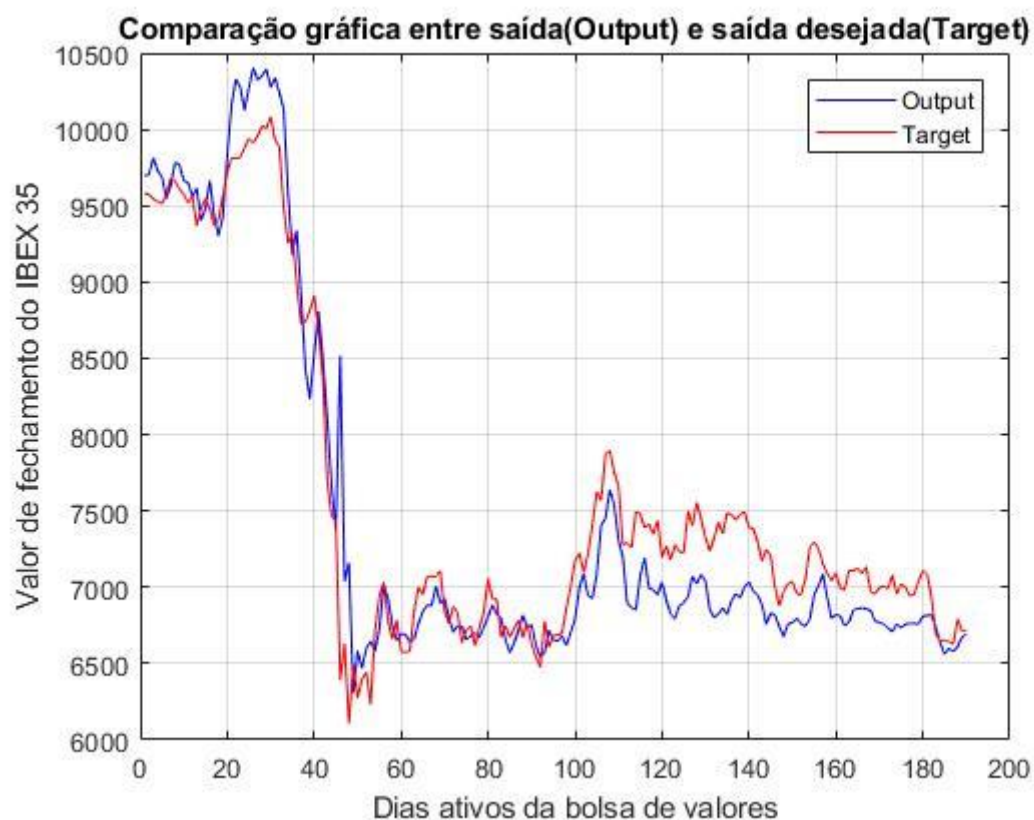
Fonte: Elaboração própria

Tabela 12 - Resultados do *MAPE* da rede neural *Feedforward* para o segundo caso com a configuração 90% treinamento, relativo à predição dos valores de fechamento do ano 2020.

FEEDFORWARD																
INPUT:SELEÇÃO DE ÍNDICES																
90%																
1 camada oculta						3 camadas ocultas										
NEURÔNIOS POR CAMADA						NEURÔNIOS POR CAMADA										
100	150	200	250	300	400	48-36-24	50-55-60	40-35-30	20-15-5	70-40-10	25-15-5	30-25-20	30-20-30	10-40-70	30-35-40	20-25-30
14.975	47.479	32.418	34.420	151.947	88.224	22.824	18.583	18.581	11.553	20.122	3.376	10.914	17.907	8.217	10.433	15.199

Fonte: Elaboração própria

Figura 17 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (25,15,5) neurônios, relativo à predição dos valores de fechamento do ano 2020.



Fonte: Elaboração própria

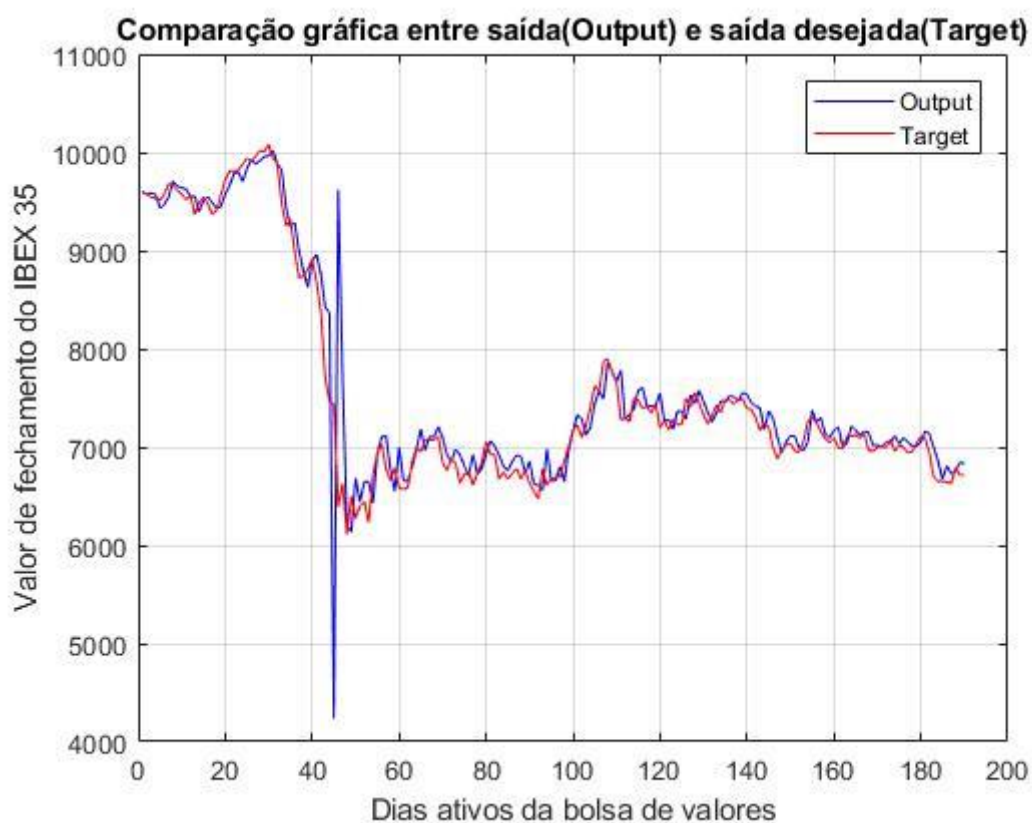
5.2.2 USANDO A REDE NEURAL CASCADE-FORWARD

Tabela 13 - Resultados do *MAPE* da rede neural *Cascade-Forward* para o segundo caso com a configuração 70% treinamento, relativo à predição dos valores de fechamento do ano 2020.

CASCADE																
INPUT:SELEÇÃO DE ÍNDICES																
70%																
1 camada oculta						3 camadas ocultas										
NEURÔNIOS POR CAMADA						NEURÔNIOS POR CAMADA										
100	150	200	250	300	400	48-36-24	50-55-60	40-35-30	20-15-5	70-40-10	25-15-5	30-25-20	30-20-30	10-40-70	30-35-40	20-25-30
55.703	29.725	146.981	131.202	147.443	86.384	70.441	56.153	40.595	2.265	31.756	26.158	75.009	22.093	55.155	35.588	16.855
																9.492

Fonte: Elaboração própria

Figura 18 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (20,15,5) neurônios, relativo à predição dos valores de fechamento do ano 2020.



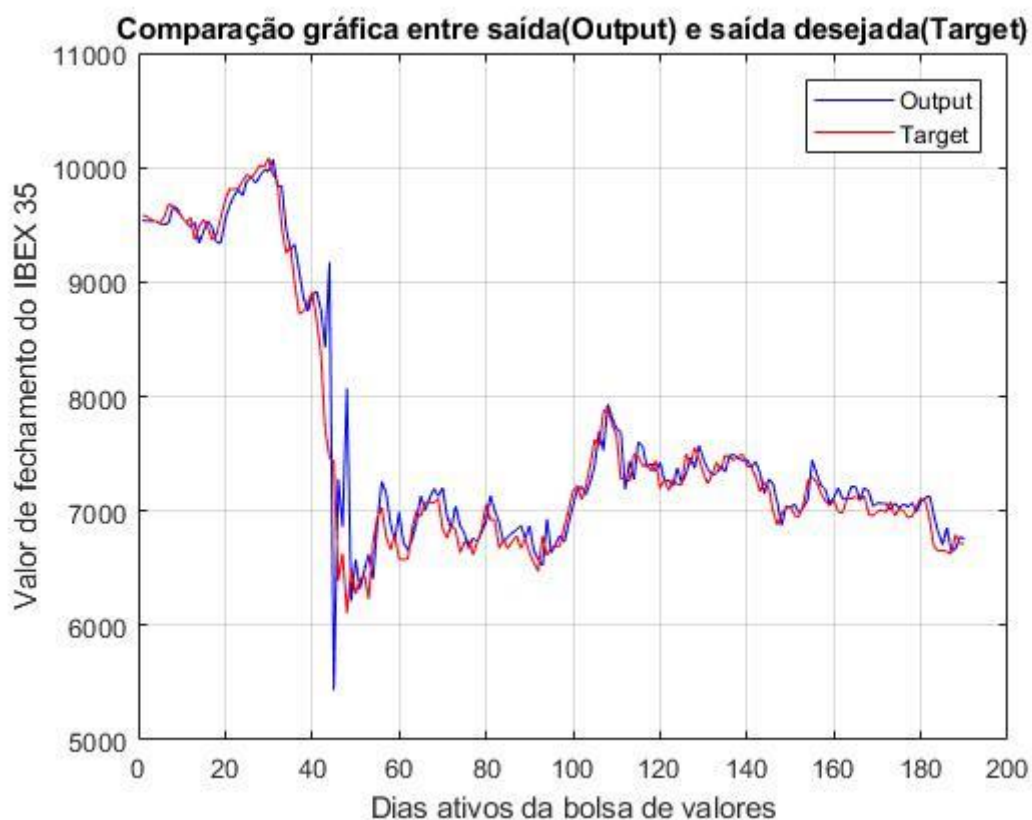
Fonte: Elaboração própria

Tabela 14 - Resultados do *MAPE* da rede neural *Cascade-Forward* para o segundo caso com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

CASCADE															
INPUT:SELEÇÃO DE ÍNDICES															
80%															
1 camada oculta						3 camadas ocultas									
NEURÔNIOS POR CAMADA						NEURÔNIOS POR CAMADA									
100	150	200	250	300	400	48-36-24	50-55-60	40-35-30	20-15-5	70-40-10	25-15-5	30-25-20	30-20-30	10-40-70	30-35-40
2.339	55.813	59.769	16.786	121.992	51.900	80.957	67.642	5.726	4.433	50.911	16.089	31.711	9.125	36.588	38.766
															2.096

Fonte: Elaboração própria

Figura 19 - Gráfico comparativo entre a previsão e o alvo correspondente a uma configuração de três camadas ocultas com (5,15,25) neurônios, relativo à predição dos valores de fechamento do ano 2020.



Fonte: Elaboração própria

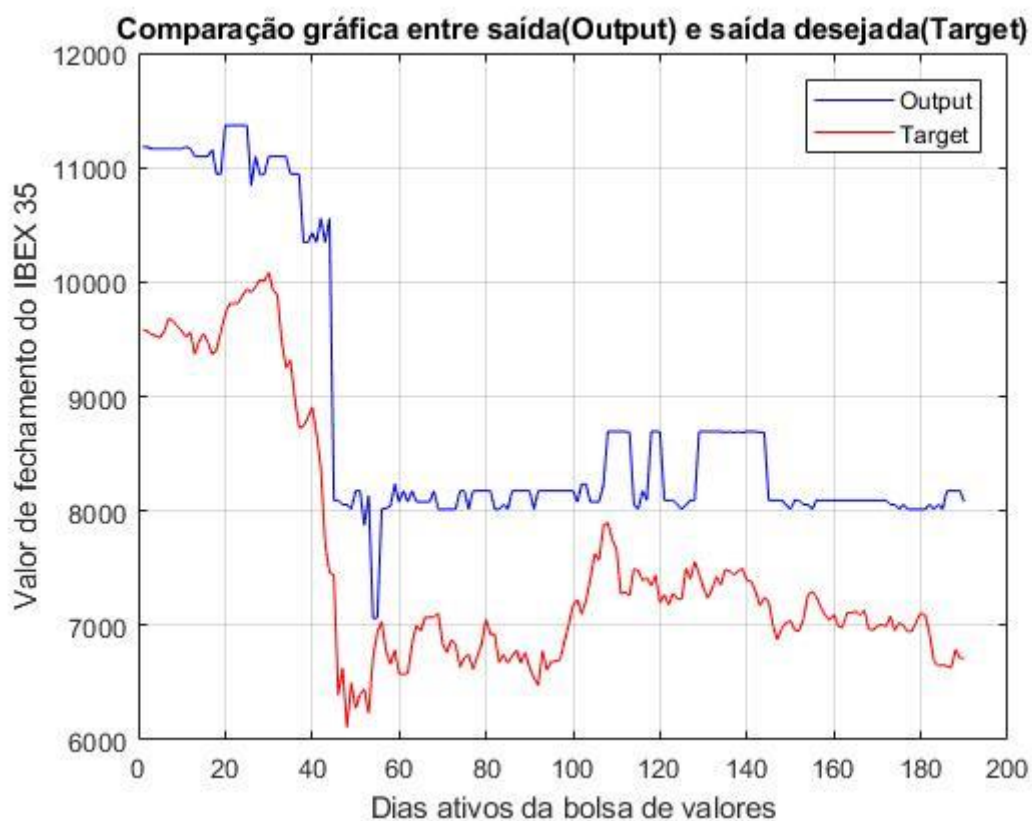
5.2.3 USANDO A GENERALIZED REGRESION NEURAL NETWORK

Tabela 16 - Resultados do *MAPE* da rede neural *GRNN* para o segundo caso com a configuração 70% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT:SELEÇÃO DE ÍNDICES	
70%	
SPREAD	
0.1	1
16.898	16.898

Fonte: Elaboração própria

Figura 21 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



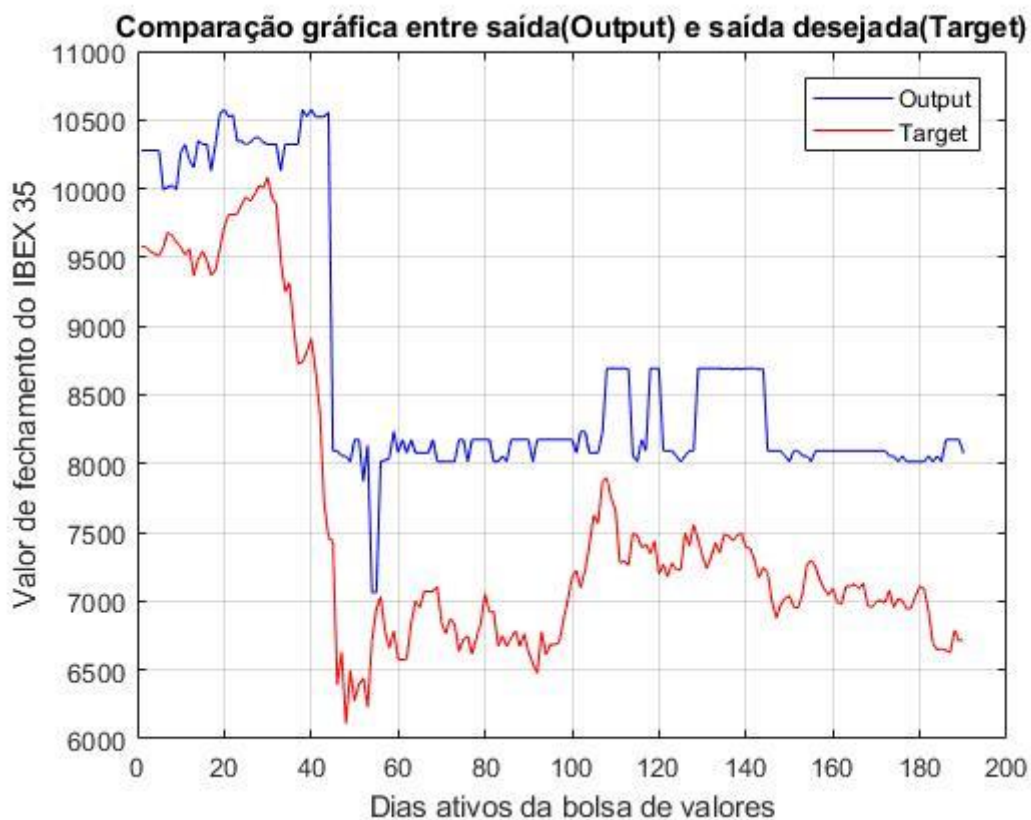
Fonte: Elaboração própria

Tabela 17 - Resultados do *MAPE* da rede neural *GRNN* para o segundo caso com a configuração 80% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT:SELEÇÃO DE ÍNDICES	
80%	
SPREAD	
0.1	1
15.254	15.254

Fonte: Elaboração própria

Figura 22 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



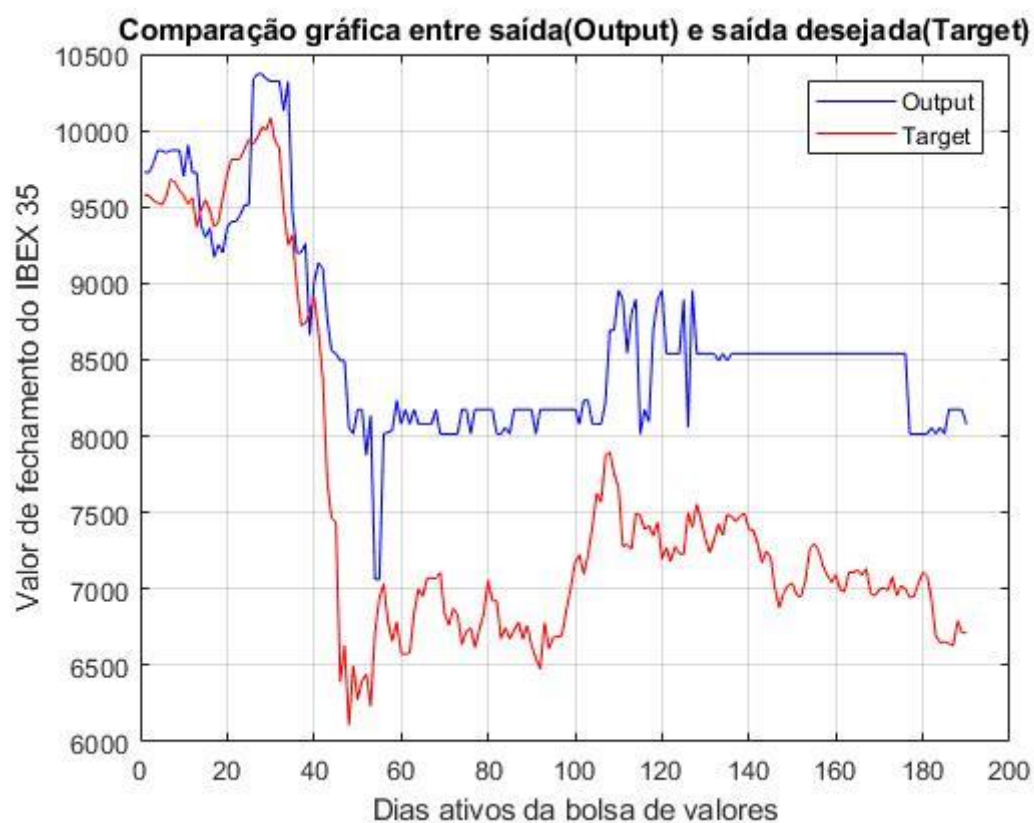
Fonte: Elaboração própria

Tabela 18 - Resultados do *MAPE* da rede neural *GRNN* para o segundo caso com a configuração 90% treinamento, relativo à predição dos valores de fechamento do ano 2020.

GRNN	
INPUT:SELEÇÃO DE ÍNDICES	
90%	
SPREAD	
0.1	1
15.243	15.243

Fonte: Elaboração própria

Figura 23 - Gráfico comparativo entre a previsão e o alvo correspondente ao *spread* igual a 1, relativo à predição dos valores de fechamento do ano 2020.



Fonte: Elaboração própria

6 CONCLUSÕES

Observa-se que os resultados mostrados do *MAPE* no caso 1 (IBEX 35 como entrada em janelamento), são melhores que no caso 2 (IBEX 35 mais índices externos como entrada em janelamento), assim como as comparações gráficas mostradas mostram um melhor desempenho da previsão realizada quando é somente utilizado o mesmo IBEX 35 em janelamento como entrada.

Estes são, os melhores resultados do *MAPE* caso 1 em cada rede neural:

Feedforward: 1.593.

Cascade-Forward: 1.647.

GRNN: 2.660.

Estes são, os melhores resultados do *MAPE* caso 2 em cada rede neural:

Feedforward: 3.376.

Cascade-Forward: 1.877.

GRNN: 15.243.

Como já foi mencionado, os resultados do caso 1 são melhores do que no caso 2, a diferença é especialmente grande quando o *GRNN* é utilizado. Além do fato de os melhores resultados do caso 1 serem melhores do que os melhores resultados do caso 2, os outros resultados correspondentes a outras configurações de neurónios mostrados nas tabelas de resultados do *MAPE* são melhores também no caso 1, mostrando uma maior fiabilidade e homogeneidade, independentemente da configuração de neurónios utilizada.

Portanto, utilizando apenas os 5 dias anteriores do mesmo IBEX 35 (janelamento) como entrada seria suficiente para fazer uma previsão muito precisa, sem a necessidade de recolher uma enorme quantidade de dados de outros índices externos, tornando a nossa tarefa de recolha de dados e condicionamento muito mais fácil. A inserção dos índices externos não só não oferece nenhuma melhoria, também pioram a previsão, especialmente ao usar a *GRNN*.

Os dados diários das entradas utilizadas para o treinamento das redes neurais foram tomados como mencionado acima de janeiro de 2010 a dezembro de 2019, e com esses dados as redes foram capazes de fazer uma previsão de excelente qualidade dos valores de fechamento da bolsa IBEX 35 do ano 2020, apesar de em 2020 termos sofrido uma pandemia que levou a uma grande crise económica.

Este artigo, expõe uma vez mais a capacidade das redes neurais de prever com precisão os índices bolsistas, quando dispomos de grandes bases de dados fiáveis sobre eles.

REFERÊNCIAS

ABOUELDAHAB, T.; FAKHRELDIN, M. **Prediction of stock market indices using hybrid genetic algorithm/ particle swarm optimization with perturbation term.**

2011. Disponível em:

<https://www.yumpu.com/en/document/view/28272800/prediction-of-stock-market-indices-using-hybrid-genetic-algorithm>. Acesso em: 18 maio 2021.

AL-DOSARY, N. M. N.; AL-HAMED, S. A.; ABOUKARIMA, A. M. K-nearest neighbours' method for prediction of fuel consumption in tractor-chisel plow systems.

Engenharia Agrícola, Jaboticabal, v. 39, n. 6, p. 729-736. Disponível em:

https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162019000600729.

Acesso em: 18 maio 2021.

ALONSO, J.; GARCIA, J. **¿Qué tan buenos son los patrones del IGBC para predecir su comportamiento?** 2009. Disponível em:

https://www.researchgate.net/publication/41805226_Que_tan_buenos_son_los_patrones_del_IGBC_para_predecir_su_comportamiento_una_aplicacion_de_datos_de_alta_frecuencia. Acesso em: 18 maio 2021.

ASADI, S.; HADAVANDI, E.; MEHMANPAZIR, F.; NAKHOSTIN, M.

Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction. **Knowledge-Based Systems**, Amsterdam, v. 35, p. 245-58, 2012.

BHARDWAJ, G.; SWANSON, N. An empirical investigation of the usefulness of Arfima models for predicting macroeconomic and financial time series. **Journal of Econometrics**, Amsterdam, v. 131, n.1/2, p. 539-578, 2006.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, Amsterdam, v. 31, n. 3, p. 307-327, 1986.

BOYACIOGLU, M. A.; AVCI, D. An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange. **Expert Systems with Applications**, Doetinchem, v. 37, n. 12, 7908-7912, 2010.

CAI, Y.; CHOU, R.; LI, D. Explaining international stock correlations with CPI fluctuations and market volatility. **Journal of Banking & Finance**, Amsterdam, v. 33, n. 11, p. 2026-2035, 2009.

CALDAS MONTES, G.; PIRES TIBERTO, B. Macroeconomic environment, country risk and stock market performance: evidence for Brazil. **Economic Modelling**, Amsterdam, v. 29, n. 5, p. 1666-1678, 2012.

CARVALHO, J. V.; ROMANEL, C. Redes neurais temporais aplicadas ao monitoramento de barragens. **Revista Eletrônica de Sistemas de Informação**, Curitiba, v. 6, n. 1, 2007. Disponível em:

<http://www.periodicosibepes.org.br/index.php/reinfo/article/view/186>. Acesso em: 18 maio 2021.

CECILIA-GARCÍA, M.; JALAL, A. M.; GARZÓN, L. A.; LÓPEZ, J. M. Métodos para predecir índices Bursátiles. **Ecos de Economía**, Medellín, v. 17, n. 37, p. 51-82, 2013. Disponível em: http://www.scielo.org.co/scielo.php?pid=S1657-42062013000200003&script=sci_abstract&tlng=es. Acesso em: 18 maio 2021.

CHAIGUSIN, S.; CHIRATHAMJAREE, C.; Y CLAYDEN, J. The use of neural networks in the prediction of the stock exchange of Thailand (SET) index. *In: IEEE COMPUTATIONAL INTELLIGENCE FOR MODELLING CONTROL & AUTOMATION*, 2008, Vienna. **Proceedings** [of the...]. Vienna: IEEE, 2008.

CHANG, Y.; YEUNG, C.; YIP, C. **Analysis of the influence of economic indicators on stock prices using multiple regression**. 2000. Disponível em: <https://silo.tips/download/analysis-of-the-influence-of-economic-indicators-on-stock-prices-using-multiple>. Acesso em: 18 maio 2021.

CHEN, A. S.; LEUNG, M.; DAOUK, H. Application of neural networks to an emerging financial market: forecasting and trading the Taiwan stock index. **Computers & Operations Research**, Oxford, v. 30, n. 6, p. 901-923, 2003.

CHEN, S. S. Predicting the bear stock market: macroeconomic variables as leading indicators. **Journal of Banking & Finance**, Amsterdam, v. 33, n. 2, p. 211-223, 2009.

CHENG, C. H.; CHEN, T. L.; WEI, L. Y. A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. **Information Sciences**, Philadelphia, v. 180, n. 9, p. 1610-1629, 2010.

CHUNG-HO, S.; TAI-LIANG, C.; CHENG, C.; CHING, C. Forecasting the stock market with linguistic rules generated from the minimize entropy principle and the cumulative probability distribution approaches. **Entropy**, Basel, v. 12, n. 12, 2010. Disponível em: https://www.researchgate.net/publication/49588637_Forecasting_the_Stock_Market_with_Linguistic_Rules_Generated_from_the_Minimize_Entropy_Principle_and_the_Cumulative_Probability_Distribution_Approaches. Acesso em: 18 maio 2021.

CLEMENTS, M.; FRANCES, P.; SWANSON, N. Forecasting economic and financial time-series with non-linear models. **International Journal of Forecasting**, Amsterdam, v. 20, n. 2, p. 169-183, 2004.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, New York, v. 20, n. 3, p. 273-297, 1995.

DAI, W.; WU, J. Y.; LU, C. J. Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. **Expert Systems with Applications**, Oxford, v. 39, n. 4, p. 4444-4452, 2012.
EL-HENAWY, I.; KAMAL, A.; ABDELBARY, H.; ABAS, A. Predicting stock index using neural network combined with evolutionary computation methods. *In: THE*

INTERNATIONAL CONFERENCE ON INFORMATICS AND SYSTEMS – INFOS, 7, 2010, Cairo. **Proceedings** [of the...]. Cairo: IEEE, 2010. 2010. 6 p.

ENGLE, R. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, West Sussex, v. 50, n. 4, p. 987-1007, 1982.

FALAHIAN, R.; DASTJERDI, M. M.; GHARIBZADEH, S. Authentic modelling of complex dynamics of biological systems by the manipulation of artificial intelligence. *In: THE INTERNATIONAL SYMPOSIUM ON ARTIFICIAL INTELLIGENCE AND SIGNAL PROCESSING, AISP, 2015, Mashhad. Proceedings* [of the...] Mashhad: IEEE, 2015. p. 47-52.

FERNÁNDEZ-BENLLOCH, L. **Predicción del índice bursátil Ibex 35 mediante selección de inputs con redes neuronales**. 2016. Disponível em: <https://riunet.upv.es/handle/10251/72570>. Acesso em: 19 maio 2021.

FRANSES, P.; GHIJSELS, H. Additive outliers, GARCH and forecasting volatility. **International Journal of Forecasting**, Amsterdam, v. 15, n. 1, p. 1-9. 1999.

GENÇTÜRK, M.; ÇELİK, I.; BINICI, Ö. Causal relations among stock returns and macroeconomic variables in a small and open economy. **African Journal of Business Management**, Durbanville, v. 9, n. 1, p. 6177-6182, 2012.

GILLILAND, M. **The business forecasting deal: exposing myths, eliminating bad practices, providing practical solutions**. New York: John Wiley, 2010. 252 p.

GURESEN, E.; KAYAKUTLU, G.; DAIM, T. Using artificial neural network models in stock market index prediction. **Expert Systems with Applications**, Oxford, v. 38, n. 8, p. 10389-10397, 2011.

GUEVARA-DÍAZ, J **Mapas autoorganizativos y el problema del agente viajero, backpropagation y reconocimiento de dígitos**. 2007. Disponível em: <http://www.vision.ime.usp.br/~jorjasso/files/SOMBACKJorgeGuevara.pdf>. Acesso em: 19 maio 2021.

GOUK, H. G.; BLAKE. A. M. Fast sliding window classification with convolutional neural networks. *In: PROCEEDINGS INTERNATIONAL CONFERENCE ON IMAGE AND VISION COMPUTING, 29, 2014, Hamilton. Proceedings* [of the...]. Hamilton: IEEE, 2014. p. 114–118.

GROSSBERG, S. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world. **Neural Networks**, Oxford, 2013. Disponível em:

https://www.researchgate.net/publication/233410858_Adaptive_Resonance_Theory_How_a_brain_learns_to_consciously_attend_learn_and_recognize_a_changing_world_Neural. Acesso em: 19 maio 2021.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. New Jersey: Prentice-Hall, 2008.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. H.; DE-JESÚS, O. **Neural network design**. 2. ed. Amsterdam: M. Hagan, 2014.

HOLLAND, J. Genetic algorithms. **Scientific American**, New York, v. 267, n. 1, p. 66-73, 1992.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. *In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON NEURAL NETWORKS, ICNN'95*, 1995, Perth. **Proceedings** [of the...]. Perth: [s.n.], 1995. V. 4, p. 1942-1948.

KIM, K. J.; HAN, I. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. **Expert Systems with Applications**, Oxford, v. 19, n. 2, p. 125-132, 2000.

KIRKPATRICK, S.; GELATT, C.; VECCHI, M. **Optimization by simulated annealing**. New York: Science, 1983.

KOMO, D.; CHANG, C. I.; KO, H. Neural network technology for stock market index prediction. *In: INTERNATIONAL SYMPOSIUM ON SPEECH, IMAGE PROCESSING AND NEURAL NETWORKS*, 1994, Hong Kong. **Proceedings** [of the...]. Hong Kong: IEEE, 1994. p. 543-546.

LASFER, M.; MELNIK, A.; THOMAS, D. Short-term reaction of stock markets in stressful circumstances. **Journal of Banking & Finance**, Amsterdam, v. 27, n. 10, p. 1959-1977, 2003.

LEE, K.; JO, G. Expert system for predicting stock market timing using a candlestick chart. **Expert Systems with Applications**, Oxford, v. 16, n. 4, p. 357-364, 1999.

LEVENBERG, K. A Method for the solution of certain non-linear problems in least squares. **Quarterly of Applied Mathematics**, Providence, v. 2, n. 2, p. 164-168, 1944. doi:10.1090/qam/10666.

LIU, H. C.; HUNG, J. C. Forecasting S&P-100 stock index volatility: the role of volatility asymmetry and distributional assumption in Garch models. **Expert Systems with Applications**, Oxford, v. 37, n. 7, p. 4928-4934, 2010.

LU, C. J.; CHANG, C. H.; CHEN, C. Y.; CHIU, C. C.; LEE, T. S. Stock index prediction: a comparison of MARS, BPN and SVR in an emerging market. *In: INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND MANAGEMENT*, 2009, Hong Kong. **Proceedings** [of the...]. Hong Kong: IEEE, 2009. p. 2343-2347.

LUPU, C.; POPESCU, D.; UDREA, A. Real-time control applications for nonlinear processes based on adaptive control and the static characteristic. **WSEAS Transactions on Systems and Control**, Zographou, v. 3, p. 607-616, 2008.

MACIÁN-PÉREZ, L. **Estudio de los efectos de las variables macroeconómicas sobre la bolsa**. 2017. Disponível em:

<https://riunet.upv.es/bitstream/handle/10251/80051/MACI%C3%81N%20-%20Estudio%20de%20los%20efectos%20de%20las%20principales%20variables%20macroecon%C3%B3micas%20sobre%20la%20Bolsa.pdf?sequence=1&isAllowed=y>. Acesso em: 19 maio 2021.

MAJHI, R.; PANDA, G.; MAJHI, B.; SAHOO, G. Efficient prediction of stock market indices using adaptive bacterial foraging optimization (ABFO) and BFO based techniques. **Expert Systems with Applications**, Oxford, v. 36, n. 6, p. 10097-100104, 2009.

MARQUARDT, D. W. An algorithm for least-squares estimation of nonlinear parameters. **SIAM Journal on Applied Mathematics**, Philadelphia, v. 11, n. 2, p. 431–441, 1963.

MORRIS, G. L. **Candlestick charting explained**: timeless techniques for trading stocks and futures. New York: McGraw-Hill, 2006.

MULLER, S.; MARCHETTO, J.; AIRAGHI, S.; KOUMOUTSAKOS, P. Optimization based on bacterial chemotaxis. **IEEE Transactions on Evolutionary Computation**, Piscataway, v. 6, n. 1, p. 16-29, 2002.

NIU, D. X.; WANG, H. Q.; GU, Z. H. Short-term load forecasting using general regression neural network. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING AND CYBERNETICS, 2005, Guangzhou. **Proceedings** [of the ...]. New York: IEEE, 2005. v. 7, p. 4076-4082, 2005.

NOSE FILHO, K. **Previsão de carga multinodal utilizando redes neurais de regressão generalizada**. 2011. 90 f. Dissertação (Mestrado) – Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2011. Disponível em: <https://repositorio.unesp.br/handle/11449/87097>. Acesso em: 19 maio 2021.

ORR, M. J. L. **Introduction to radial basis function networks**. Edinburgh: Centre for Cognitive Science, University of Edinburgh, 1996. Disponível em: <https://www.cc.gatech.edu/~isbell/tutorials/rbf-intro.pdf>. Acesso em: 19 maio 2021.

PAI, P. F.; LIN, C. S. A hybrid Arima and support vector machines model in stock price forecasting. **Omega**, Oxford, v. 33, n. 6, p. 497-505, 2005.

PARISI, A.; PARISI, F.; DÍAZ, D. Modelos de algoritmo genético y redes neuronales en la predicción de índices bursátiles asiáticos. **Cuadernos de Economía**, Santiago, v. 43, n. 128, p. 251-284, 2006. Disponível em: https://scielo.conicyt.cl/scielo.php?script=sci_abstract&pid=S0717-68212006000200002&lng=en&nrm=iso. Acesso em: 19 maio 2021.

PARK, D. C.; EL-SHARKAWI, M. A.; MARKS II, R. J.; ATLAS, E. Electric load forecasting using an artificial neural network. **IEEE Transactions on Power Systems**, Piscataway, v. 6, n. 2, p. 442-449, 1991.

PIERDZIOCH, C.; DOPKE, J.; HARTMANN, D. Forecasting stock market volatility with macroeconomic variables in real time. **Journal of Economics and Business**, Philadelphia, v. 60, n. 3, p. 256-276, 2008.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Cornell Aeronautical Laboratory, Psychological Review**, Whashington, v. 65, n. 6, p. 386–408, 1958. Disponível em: <https://psycnet.apa.org/fulltext/1959-09865-001.pdf>. Acesso em: 19 maio 2021.

ROH, T. Forecasting the volatility of stock price index. **Expert Systems With Application**, Oxford, v. 33, n. 4, p. 916-922, 2007.

SHEN, W.; GUO, X.; WUB, C.; WU, D. Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. **Knowledge-Based Systems**, Amsterdam, v. 24, n. 3, p. 378-385, 2011.

SILVERMAN, B.; JONES, M. An important contribution to nonparametric discriminant analysis and density estimation: commentary on fix and hedges (1951). **International Statistical Review / Revue Internationale De Statistique**, West Sussex, v. 57, n. 3, p. 233-238, 1989.

SPECHT, D. F. A generalized regression neural network. **IEEE Transactions on Neural Networks**, Piscataway, v. 2, p. 568-576, 1991.

STIGLER, S. Who discovered bayes's theorem? **The American Statistician**, Alexandria, v. 37, n. 4, p. 290-296, 1983.

TORO-OCAMPO, E. M.; MOLINA-CABRERA, A.; GARCÉS-RUIZ, A. Pronóstico de bolsa de valores empleando técnicas inteligentes. **Tecnura**, Bogota, v. 9, n. 18, p. 57-66, 2006.

TRIPATHI, M. M.; UPADHYAY, K. G.; SINGH, S. N. Short-term load forecasting using generalized regression and probabilistic neural networks in the electricity market. **The Electricity Journal**, Oxford, v. 21, n. 9, p. 24-34, 2008.

VELÁSQUEZ-HENAO, J. D.; FRANCO-CARDONA, C. J.; OLAYA-MORALES, Y. A review of DAN2 (dynamic architecture for artificial neural networks) model in time series forecasting. **Ingeniería y Universidad**, Bogotá, v. 16, n. 1, p. 135-146, 2012. Disponível em: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-21262012000100008. Acesso em: 19 maio 2021.

WANG, J. Z.; WANG, J. J.; ZHANG, Z. G.; GUO, S. P. Forecasting stock indices with back propagation neural network. **Expert Systems with Applications**, Oxford, v. 38, n. 11, p. 14346-14355, 2011.

WANG, J. J.; WANG, J. Z.; ZHANG, Z. G.; GUO, S. P. Stock index forecasting based on a hybrid model. **Omega**, Oxford, v. 40, n. 6, p. 758-766, 2012.

WANG, Y. F.; CHENG, S.; HSU, M. H. Incorporating the Markov chain concept into fuzzy stochastic prediction of stock indexes. **Applied Soft Computing**, Amsterdam, v. 10, n. 2, p. 613-617, 2010.

WARSITO, B.; SANTOSO, R.; SUPARTI, S.; YASIN, H. Cascade forward neural network for time series prediction. **Journal of Physics: Conference Series**, Bristol, v. 1025, n. 1, p. 012097, 2018. Disponível em: https://www.researchgate.net/publication/325445281_Cascade_Forward_Neural_Network_for_Time_Series_Prediction. Acesso em: 19 maio 2021.

WERBOS, P. J. **Beyond regression**: new tools for prediction and analysis in the behavioral sciences. 1974. Thesis (Doctor) - Harvard University, Harvard, 1974.

XIA, C.; WNAG, J.; MCMENEMY, K. Short-medium- and long-term load forecasting model and virtual load forecaster based on radial basis function neural networks. **Electrical Power and Energy Systems**, New York, v. 32, p. 743-750, 2010.

XIN, Y. **Linear regression analysis**: theory and computing. Sigapura: World Scientific, 2009. 328 p.

YAP, K. S. Short term load forecasting using a hybrid neural network. *In*: IEEE INTERNATIONAL POWER AND ENERGY CONFERENCE - PECON, 2006, Tehran. **Proceedings** [of the ...] Tehran: IEEE, 2006. p. 123-128.

YU, T. K.; HUARNG, K. H. A neural network-based fuzzy time series model to improve forecasting. **Expert Systems With Applications**, Oxford, v. 37, n. 4, p. 3366-3372, 2010.

YUDONG, Z.; LENAN, W. Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. **Expert Systems with Applications**, Oxford, v. 36, n. 5, p. 8849-8854, 2009.

ZEMKE, S. Non-linear index prediction. **Physica A: Statistical Mechanics and its Applications**, Amsterdam, v. 269, n. 1, p. 177-183, 1999.