

Previsão do índice de ações IBEX 35 usando redes neurais artificiais

Salvador Falcón Canillas¹

FEIS/UNESP, Ilha Solteira, SP

Resumo. A previsão de índices de ações 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. 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 utilizar são *Feedforward*, *Cascade-Forward* e *Generalized Regression*.

Palavras-chave. previsão, ibex 35, bolsa de valores, backpropagation, feedforward, cascade, regressão generalizada.

1 Introdução

Hoje em dia, a previsão dos índices de ações das diferentes bolsas internacionais é uma questão importante na economia global, uma vez que movimenta enormes quantidades de dinheiro.

Os fatores que influenciam os índices de ações 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. [4]

Também, os fatores que podem influenciar um país são os índices de ações de outros países, ou fatores como 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.

Nesta pesquisa, foi realizada, uma previsão do índice IBEX 35, usando a técnica de janelamento [2]. É realizada a previsão empregando-se três redes neurais diferentes, comparando o desempenho de cada uma delas em dois casos diferentes, onde outros índices de ações, matérias-primas e taxas de câmbio tomadas como entrada:

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

As redes neurais a serem utilizadas são *Feedforward*, *Cascade-Forward*, e Regressão Generalizada (*GRNN*).

¹sf.canillas@unesp.br

- 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)

2 Redes neurais utilizadas

2.1 *Feedforward neural network (FFNN)*

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. Neste caso, o algoritmo utilizado seria o algoritmo backpropagation, onde a sinal resultante da saída é comparado com a saída desejada e a margem de erro de cada uma das saídas é calculada. A retropropagação (*backpropagation*) propaga para trás as saídas de erro para todos os neurônios da camada oculta que contribuem para a saída. Entretanto, os neurônios da camada oculta, recebe apenas uma fração do sinal de erro total, baseado na contribuição relativa de cada neurônio para a saída original. Isto é feito repetidamente até que os neurônios da rede tenham recebido um sinal de erro que descreva sua contribuição relativa para o erro total. Os pesos de conexão de cada neurônio são então atualizados para obter um resultado e para classificar corretamente os padrões de treinamento. [1]

2.2 *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.

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. [11]

2.3 *Generalized regression neural network (GRNN)*

GRNN, proposta por Specht (1991) [9], é uma rede capaz de estimar uma variável dependente de uma variável independente por meio 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. Possui uma estrutura fixa, dada por uma camada de entrada, duas camadas intermediárias 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. [6].

Esta rede tem sido amplamente utilizada nos últimos anos, principalmente para problemas relacionados com a previsão de séries temporais, tais como taxas de eletricidade de curto prazo [6], [7], [10], [13], [14].

A *GRNN* convencional foi programada com o auxílio das *toolboxes* das redes neurais artificiais do *MATLAB* por meio da função *newgrnn*. A regressão consiste em estimar uma variável dependente a partir de uma variável independente usando-se um conjunto finito de variáveis. Na teoria dos sistemas, a variável dependente corresponde à saída e a variável independente corresponde à entrada e pode ser descrita por meio de vetores, cujo número de elementos indica o número de entradas e saídas do sistema.

3 Metodologia

Para calcular o valor de fechamento do dia seguinte, foi utilizada a conhecida técnica do janelamento [2], que consiste em inserir os valores de entrada dos dias anteriores ao valor que se precisa saber. Neste caso, entrariam-se 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.

3.1 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 princípio, os dados históricos foram tomados de janeiro de 2010 a dezembro de 2019 dos mais importantes índices de ações 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 de 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), SP ASX 200 (Austrália) SP TSX 60 (Canadá), SMI 30 (Suíça), Taiwan Weighted (Taiwan) e o próprio IBEX 35 (Espanha). Os 3 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 nossas redes neurais treinadas com os dados de 2010 a 2019 seriam capazes de prever com precisão os valores de fechamento do ano 2020, que como todos sabemos foi um ano em que sofremos uma pandemia global e, portanto, o mercado de ações sofreu quedas significativas.

3.2 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 usados como entrada, apenas os dados de 5 dias anteriores ao valor do dia do mesmo índice IBEX 35 que se quer 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 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.

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 3 redes neurais diferentes.

4 Resultados obtidos

A seguir, são apresentados os resultados (verificação e testes) obtidos, via o emprego das RNA *Feedforward* e *Cascade-Forward*, variando-se o número de camadas ocultas, o número de neurônio em cada camada oculta, assim como o percentual da base de dados usado na fase de treinamento. Nos casos do uso das RNA *Feedforward* e *Cascade-Forward* foram empregados o algoritmo de treinamento *Levenberg-Marquardt Backpropagation* [5].

A função de transferência usada é a sigmoide tangente hiperbólica. Os parâmetros usados são: (1) 90% e 10% da base de dados para as fases de treinamento e de testes, respectivamente; (2) número de camadas ocultas: 1 e 3. No caso RNA *GRNN*, o único parâmetro mudado é a dispersão (*spread*), em razão do número de neurônios da camada oculta que deve ser igual ao número de amostras do treinamento. Usa-se, neste caso, o $spread = 1$. Neste artigo, são apresentados os gráficos (figuras) comparativos do comportamento da saída obtida pela RNA (azul) em relação à curva real (target) (vermelho) durante a fase de testes. A qualidade dos resultados obtidos é inferida mediante a análise da métrica mean absolute percentage error (*MAPE*) [8] assim definido:

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

onde:

MAPE : Mean Absolute Percentage error;

NA : Número de 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 ;

As tabelas de *MAPE* para cada rede neural de ambos casos também são mostradas, onde a célula marcada na cor vermelho mostra a configuração que resultou com um valor de *MAPE* mais baixo. Portanto, a configuração que gerou o *MAPE* mais baixo foi representada graficamente neste artigo.

Os resultados obtidos referem-se ao fechamento do índice IBEX 35 relativo ao período compreendido entre 01/01/2020 e 30/09/2020. Ressalta-se a dificuldade de realizar a previsão por ser um ano atípico, por se tratar da ocorrência da pandemia da COVID-19. No referido ano as bolsas de valores tiveram muitas flutuações por causa da crise de amplitude mundial.

4.1 Análise Gráfica do caso 1

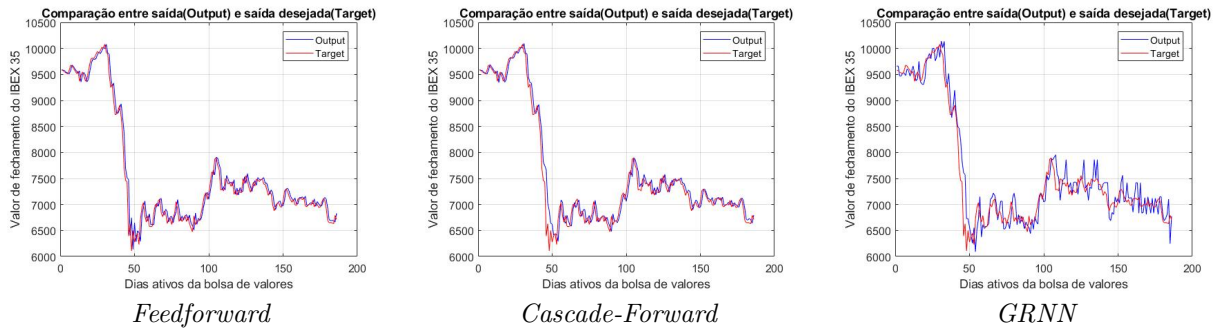


Tabela 1: Gráficos comparativos entre as saídas e o alvos correspondentes ao caso 1, referentes à predição dos valores de fechamento do ano 2020.

4.2 Resultados de MAPE do caso 1

| FEEDFORWARD | | | | | | | | | | | | | | | |
|----------------------|-------|-------|-------|-------|-------|-------|----------------------|-------|-------|---------|----------|-------|-------|-------|-------|
| 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 | 4-3-2 |
| 2.116 | 2.707 | 9.941 | 2.860 | 3.642 | 2.358 | 3.268 | 1.632 | 1.610 | 1.664 | 1.754 | 2.118 | 1.922 | 1.685 | 3.895 | 1.645 |

| 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 | 4-3-2 |
| 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 | 1.720 |

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

Tabela 2: Tabelas de valores do MAPE para cada rede neural do primeiro caso

4.3 Análise Gráfica do caso 2

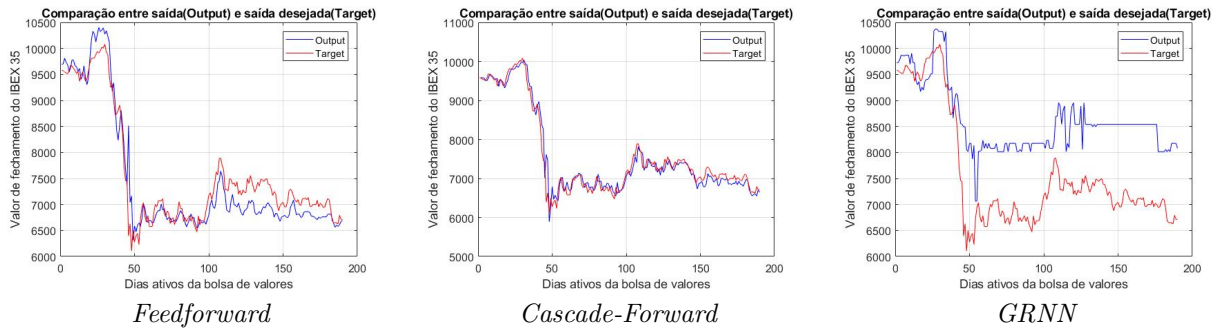


Tabela 3: Gráficos comparativos entre a saídas e o alvos correspondentes ao caso 2, referentes à predição dos valores de fechamento do ano 2020.

4.4 Resultados de MAPE do caso 2

| FEEDFORWARD | | | | | | | | | | | | | | | | | |
|--------------------------|--------|--------|---------|---------|--------|----------------------|----------|----------|---------|----------|---------|----------|----------|----------|----------|----------|---------|
| INPUT:SELEÇÃO DE ÍNDICES | | | | | | | | | | | | | | | | | |
| 90% | | | | | | | | | | | | | | | | | |
| 1 camada oculta | | | | | | 3 camadas ocultas | | | | | | | | | | | |
| NEURÓNIOS POR CAMADA | | | | | | NEURÓNIOS 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 | 5-15-25 |
| 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 | 5.258 |
| CASCADE | | | | | | | | | | | | | | | | | |
| INPUT:SELEÇÃO DE ÍNDICES | | | | | | | | | | | | | | | | | |
| 90% | | | | | | | | | | | | | | | | | |
| 1 camada oculta | | | | | | 3 camadas ocultas | | | | | | | | | | | |
| NEURÓNIOS POR CAMADA | | | | | | NEURÓNIOS 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 | 5-15-25 |
| 1.877 | 34.187 | 12.967 | 122.998 | 51.564 | 82.196 | 34.178 | 24.602 | 23.146 | 9.495 | 11.775 | 8.859 | 14.304 | 21.327 | 12.942 | 92.118 | 7.319 | 2.524 |
| GRNN | | | | | | | | | | | | | | | | | |
| INPUT:SELEÇÃO DE ÍNDICES | | | | | | | | | | | | | | | | | |
| 90% | | | | | | | | | | | | | | | | | |
| SPREAD | | | | | | | | | | | | | | | | | |
| 0.1 | | | | | | | | | | 1 | | | | | | | |
| 15.243 | | | | | | | | | | 15.243 | | | | | | | |

Tabela 4: Tabelas de valores do MAPE para cada rede neural do segundo caso

5 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 evidenciam um melhor desempenho da previsão realizada quando é somente utilizado o mesmo IBEX 35 como entrada.

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*.

Referências

- [1] Fernández-Benlloch, L. Predicción del índice bursátil Ibex 35 mediante selección de inputs con redes neuronales, 2016. <http://hdl.handle.net/10251/72570>.
- [2] Gouk, H. G. and 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.
- [3] Grossberg, S. Adaptive resonance theory: how a brain learns to consciously attend, learn, and recognize a changing world, *Neural Networks*, Vol. 37, pp.1-47, 2013. <https://doi.org/10.1016/j.neunet.2012.09.017>.
- [4] Macián-Pérez, L. Estudio de los efectos de las variables macroeconómicas sobre la bolsa. 2017. Disponível em: <https://riunet.upv.es>. Acesso em: 19 maio 2021.
- [5] Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters, *SIAM Journal on Applied Mathematics*, Vol. 11, No. 2: pp. 431–441, 1963. DOI:10.1137/0111030.
- [6] Niu, D. X. Short-term load forecasting using general regression neural network. *International conference on machine learning and cybernetics. Proceedings...New York: IEEE*, v. 7, p. 4076-4082, Guangzhou, 2005.
- [7] Nose Filho, K. Previsão de carga multinodal utilizando redes neurais de regressão generalizada. 90 f. Dissertação (mestrado), *Universidade Estadual Paulista, Faculdade de Engenharia de Ilha Solteira*, 2011.
- [8] Park, D. C., El-Sharkawi, M. A., Marks II, R. J., Atlas, E. Electric load forecasting using an artificial neural networks. *IEEE Transactions on Power Systems*, Vol. 6, No. 2, pp. 442-449, Damborg, 1991. DOI: 10.1109/59.76685.
- [9] Specht, D. F. A generalized regression neural network. *IEEE Transactions on Neural Networks*, v. 2, p. 568-576, New York, 1991.
- [10] 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*, v. 21, n. 9, p. 24-34., Oxford, 2008.
- [11] Warsito, B., Santoso, R., Suparti, S., Yasin, H. Cascade Forward Neural Network for Time Series Prediction. *Journal of Physics: Conference Series* 1025.012097., 2018.
- [12] Werbos, P. J. Beyond regression: new tools for prediction and analysis in the behavioral sciences. (*PhD. Thesis*)-*Harvard University*, Harvard, 1974.
- [13] 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*, v. 32, p. 743-750, New York, 2010.
- [14] Yap, K. S. Short term load forecasting using a hybrid neural network. *IEEE International Power and Energy Conference - PECon. .Proceedings... New York:IEEE*, p.123-128., Tehran, 2006.