

# **PREVISIBILIDADE DE MERCADOS DE CAPITAIS: UM ESTUDO COMPARATIVO ENTRE PAÍSES DESENVOLVIDOS E EM DESENVOLVIMENTO**

**EVERTON ANGER CAVALHEIRO**  
UNIVERSIDADE FEDERAL DE SANTA MARIA

**KELMARA MENDES VIEIRA**  
UNIVERSIDADE FEDERAL DE SANTA MARIA

**PAULO SERGIO CERETTA**  
UNIVERSIDADE FEDERAL DE SANTA MARIA

**LARISSA DE LIMA TRINDADE**  
CENTRO UNIVERSITÁRIO FRANCISCANO

## **Resumo**

A pergunta de pesquisa deste artigo é: “o uso de redes polinomiais GMDH é suficientemente capaz de demonstrar previsibilidade do retorno mensal dos índices dos países pertencentes ao BRIC’s e aos países pertencentes ao G4?”. Para responder esta pergunta, estabeleceu-se o seguinte objetivo: comparar a previsibilidade do mercado de capitais dos países pertencentes ao BRIC’s (Brasil, Rússia, Índia e China) e as quatro maiores economias do mundo atual (EUA, Japão, Alemanha e Reino Unido), através do uso de redes neurais Group Method of Data Handling (GMDH) em uma aproximação indutiva do retorno mensal, na forma logaritma, dos índices mais representativos destes países. Para atender este objetivo, inicialmente, foi realizada uma revisão teórica sobre o tema de redes neurais polinômiais, em especial destaque as do tipo GMDH e após, procurou-se a desmistificação do conceito de redes neurais, prejudicada pelo efeito “caixa preta” que impede o conhecimento dos métodos de processamento e decisão de uma rede neural, bem como seus efeitos. Os resultados das modelagens identificadas nas redes demonstram uma maior previsibilidade para o mercado brasileiro.

Palavras-chave: Previsão, mercado de capitais, redes neurais, séries temporais.

## **1. INTRODUÇÃO**

Durante as últimas décadas, a área da pesquisa em Redes Neurais Artificiais (RNAs) vem atraindo muita atenção e se expandindo rapidamente, com aplicações em inúmeras áreas do conhecimento, sendo usadas, conforme cita Palma Neto e Nicoletti (2005, p. 17), na solução de problemas que envolvem classificação, controle e predição.

Person (2007, p. 1), cita que o interesse pela previsão dos preços dos retornos das ações é provavelmente tão antigo quanto os próprios mercados, bem como a literatura sobre o assunto é vasta e significativa. Neste sentido, o uso de redes neurais em modelos de previsão tem se mostrado cada vez mais presente na literatura e nas modelagens empíricas nacionais e internacionais.

Ludwing Jr. e Costa (2007, p. 7) citam que uma das desvantagens das redes neurais é o fato delas, normalmente, serem uma “caixa preta”, sendo impossível saber o motivo de uma rede chegar a um resultado.

Valença (2005, p. 23) ao introduzir, no Brasil, as redes neurais *Group Method of Data Handling* (GMDH), cita que tais modelagens podem ser vistas como um sistema híbrido composto por uma combinação de redes neurais e conceitos estatísticos capazes de atenuar este efeito “caixa preta”, pela sua simplicidade de implantação. Embora sua aplicação e entendimento sejam simples, raros estudos têm sido desenvolvidos através de modelagens GMDH, desta forma, a pergunta de pesquisa deste artigo é: “o uso de redes polinomiais GMDH é suficientemente capaz demonstrar previsibilidade do retorno mensal dos índices dos países pertencentes ao BRIC’s e aos países pertencentes ao G4?”.

O objetivo geral deste trabalho é comparar a previsibilidade dos mercados de capitais dos principais países emergentes, pertencentes ao BRICs, e os quatro países mais desenvolvidos economicamente, através do uso de redes neurais GMDH para séries de dados do retorno mensal dos índices destes países. Secundariamente, procura-se desvelar o conceito de redes neurais polinomiais, pouco difundidas no Brasil, e desmitificar o conceito de “caixa preta” que prejudica o entendimento de como são tomadas as decisões durante o cálculo de uma rede neural, fato que também prejudica o entendimento das possíveis consequências do uso de tais procedimentos de cálculo.

## 2. REDES NEURAI ARTIFICIAIS

As Redes Neurais Artificiais (RNAs) são um conjunto de técnicas que tentam simular, em meio computacional, o funcionamento do cérebro humano de uma maneira simplificada. Elas são capazes de reconhecer padrões, extrair regularidades e detectar relações subjacentes em um conjunto de dados aparentemente desconexos.

Uma Rede Neural Artificial (RNA), tal como seu paralelo biológico, é composto de um certo número de nós, ou *units*, também denominados neurônios, conectados por ligações, conexões sinápticas ou links. Cada link possui uma quantidade associada a um peso. O conhecimento da rede é adquirido pela atualização destes pesos. Algumas das *units* são conectadas ao ambiente externo,

Cada neurônio tem um conjunto de entradas ligadas de outros neurônios e outro conjunto de saídas conectadas a outros neurônios. Conforme o nível atual das entradas do neurônio, conjugadas com os pesos sinápticos das conexões, o valor resultante irá ativar ou não as saídas deste neurônio. Esta conjugação pode ser representada como uma combinação das entradas ponderadas pelos pesos sinápticos, conforme demonstrado na Figura [1].

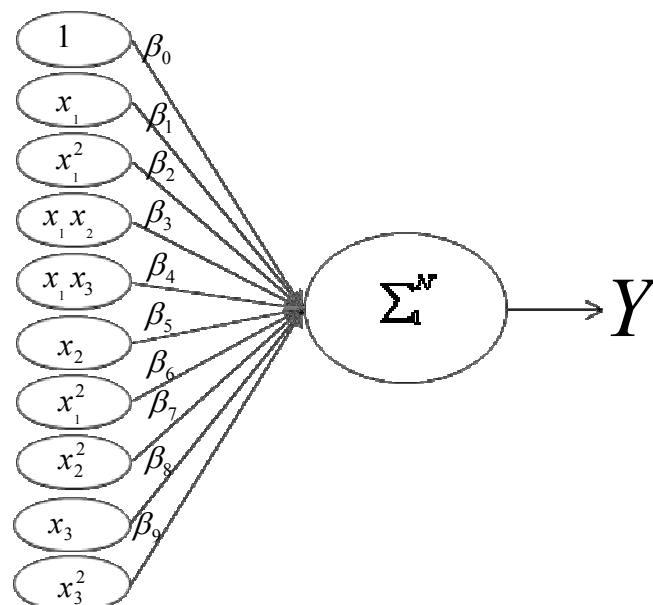


Figura 1 – Detalhe de um neurônio GMDH.

Na Figura [1], percebe-se que neste neurônio há uma combinação de *inputs* lineares e não lineares, o que torna sua aplicação no mercado financeiro muito objetiva.

Para Haykin (2001, p. 28), rede neural pode ser definida como um processador maciço e paralelamente distribuído, constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. A rede neural assemelha-se ao cérebro, em dois aspectos: (1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (2) forças de conexão entre neurônios (pesos sinápticos) são utilizados para armazenar o conhecimento adquirido

Ludwing Jr. e Costa (2007, p. vii) citam que as Redes Neurais Artificiais (RNAs) estão fundamentadas em diversas disciplinas, tais como neurociência, matemática, estatística, física, ciência da computação. O emprego destes algoritmos é igualmente amplo. Modelagem, análise de séries temporais, reconhecimento de padrões, processamento de sinais e controle, são exemplos de aplicações para estas técnicas.

A maior virtude das redes neurais é a capacidade de aprender a partir dos dados de entrada, com ou sem professor. Esta habilidade tem tornado o uso destes algoritmos cada vez mais freqüente. Desde seu surgimento, esta técnica de Inteligência Artificial (IA), foi a que mais evoluiu, despertando o interesse de um grande número de pesquisadores em todo o mundo.

As Redes Neurais (RNA) podem ser distinguidas com base na direção na qual o sinal flui. Basicamente há dois tipos de Redes Neurais Artificiais: *feedforward* e *feedback*. Nas redes *feedforward*, os sinais se propagam em apenas uma direção a partir da unidade de entrada, passando pelas camadas intermediárias até a saída. Nas redes *feedback*, os sinais de entrada podem propagar da saída de qualquer neurônio para a entrada em um outro neurônio.

As RNAs apresentam duas fases distintas em sua metodologia que são denominadas de aprendizagem e teste. Na fase de aprendizagem, são apresentados estímulos a entrada, denominados padrões de treinamento que farão com que a rede aprenda com os dados. Na fase de teste é demonstrada a capacidade de generalização da rede, pois seus resultados deverão ser significativos após seus pesos terem sido ajustados na fase anterior.

### 2.1. Redes polinomiais GMDH

No contexto estocástico da análise de dados, encontra-se a rede GMDH. As primeiras investigações foram realizadas por Alexey Ivakhnenko em 1968 e seguidas por Shankar, em 1972, o qual apresentou o algoritmo do mesmo nome como um método que permitia descobrir de forma sucessiva um sistema complexo de relações a partir de simples operações matemáticas. Trata-se de um método para solucionar problemas de estilo, identificação, predição de curto prazo de processos aleatórios e reconhecimento de padrões. O conteúdo deste algoritmo se desenvolveu como veículo para identificar relações lineares e não lineares entre *inputs e outputs*, gerando uma estrutura que tende a ótima, a partir de um processo sucessivo de várias manipulações dos dados, mediante a incorporação de novas camadas.

O modelo GMDH pode ser analisado como uma combinação de redes neurais e conceitos estocásticos, conforme cita Valença (2005, p. 24). Na sua formulação original, cada neurônio da camada escondida, a ser construída, recebe duas entradas e tem como função de ativação um polinômio de segundo grau. Portanto, através da combinação de cada par desses neurônios de entrada, será gerada uma função de saída polinomial cuja complexidade será função do número de camadas, isto é, se existir duas camadas, tem-se uma função polinomial de quarto grau, para três camadas, uma função polinomial de oitavo grau. Desta forma, tais redes são chamadas de redes polinomiais, pois o modelo que resulta é uma função polinomial.

### 2.2. Algoritmo GMDH

Algoritmo é um processo de cálculo em que se estipulam, com generalidade e sem restrições, regras formais para a obtenção do resultado ou da solução do problema. Nas redes GMDH são desenvolvidos, entre outros, o proposto por Ivakhnenko (1971, p. 364), em que os neurônios gerados nas camadas intermediárias são funções polinomiais de segundo grau compostos de duas variáveis. No caso de três entradas  $x_1$ ,  $x_2$  e  $x_3$ , há produção de saídas  $y$  que são uma combinação quadrática dessas entradas, podendo desdobrar-se da seguinte forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_1 x_2 + \beta_4 x_1 x_3 + \beta_5 x_2 x_3 + \beta_6 x_2 + \beta_7 x_2^2 + \beta_8 x_3 + \beta_9 x_3^2 + \xi. \quad [1]$$

Em [1],  $y$  é a saída estimada,  $x_1$ ,  $x_2$  e  $x_3$  as variáveis de entrada,  $\beta_i$  seus pesos e  $\xi$  o erro.

Como salienta Valença (2005, p. 26), cada camada construída terá uma quantidade de neurônios que depende do número de variáveis de entrada. Supondo que o número de variáveis de entrada é  $n$ , o número de neurônios na primeira camada será de  $C_2^n$ . Esses neurônios têm então seu desempenho avaliado através de uma função objetivo com limiar que é utilizada para avaliar o desempenho de cada neurônio e para decidir quais os neurônios deverão continuar e quais neurônios deverão parar. Este processo é então repetido até que reste apenas um neurônio escondido, ou então quando as camadas adicionais não provocarem nenhuma melhoria significativa.

Valença (2005, p. 27) salienta que um importante aspecto a ser considerado no algoritmo GMDH é o uso de um critério externo de seleção dos melhores neurônios em cada camada construída. Vários são os critérios de seleção que podem ser utilizados para decidir sobre a qualidade de cada modelo testado, de tal forma que o autor detalha alguns deles: i) regularidade, que utiliza o erro quadrático médio sobre o conjunto de validação; ii) critério de mínima tendência, que tem por objetivo minimizar a diferença entre os valores calculados para os conjuntos de dados A e B e iii) critério de balanço de variáveis, sendo este mais indicado para previsões de longo prazo.

### 2.3. Algoritmo combinatorial

O aspecto básico do algoritmo combinatorial é a sua estrutura em uma única camada e, de acordo com Valença (2005, p. 29), funções de soma são geradas para todas as combinações de variáveis de entrada de forma similar a uma busca polinomial. Esse algoritmo se torna eficiente pela utilização de uma técnica recursiva dos mínimos quadrados, uma vez que os modelos gerados são lineares em relação aos parâmetros.

Para um melhor entendimento, supondo-se uma função com três variáveis de entrada ( $x_1$ ,  $x_2$  e  $x_3$ ), e uma variável de saída, haverá uma busca polinomial do melhor modelo até o segundo grau, devendo se começar pelo modelo mais simples ( $y = \beta_0 + \xi$ ) até o mais complexo, semelhante ao demonstrado na Equação [1]. A seguir, demonstra-se o incremento gradual da complexidade para este algoritmo contendo uma variável:

$$y = \beta_0 + \xi, \quad [2]$$

$$y = \beta_1 x_1 + \xi, \quad [3]$$

$$y = \beta_2 x_1^2 + \xi, \quad [4]$$

...

$$y = \beta_9 x_3^2 + \xi. \quad [5]$$

Como demonstrado nas Equações [2, 3, 4 e 5], existirão  $C_{10}^1 = 10$  neurônios, sendo os pesos determinados pelo método dos mínimos quadrados, com um conjunto de treino. Na sequência, os erros obtidos por cada equação são comparados a um limiar estabelecido para cada função objetivo (complemento externo), utilizando-se um conjunto de dados teste. Finalmente, as unidades selecionadas, ditas sobreviventes serão consideradas aptas para terem sua complexidade evoluída.

Num segundo momento, todos os neurônios com duas variáveis são determinados, totalizando  $C_{10}^2 = 45$  equações parciais neste segundo passo:

$$y = \beta_0 + \beta_1 x_1 + \xi, \quad [6]$$

$$y = \beta_0 + \beta_2 x_1^2 + \xi, \quad [7]$$

...

$$y = \beta_8 x_3 + \beta_9 x_3^2 + \xi. \quad [8]$$

Os neurônios serão construídos até que no último passo obtenha  $C_{10}^{10} = 1$  combinações, que é o polinômio completo demonstrado na Equação [1]. É importante notar que, dependendo se o número de entradas for grande, irá ocorrer uma explosão de combinações, pois o número de modelos que devem ser testados aumenta rapidamente com o número de variáveis de entrada, como demonstra:  $M=2^n-1$ . No exemplo de três variáveis de entrada ( $x_1$ ,  $x_2$  e  $x_3$ ) acrescida de uma constante, que se desdobraram em 10 variáveis, demonstrada na Figura [1], haverá  $M=2^{10}-1 = 1.023$  neurônios totalmente diferentes.

#### 2.4. Rede *multilayer*

De acordo com Valença (2005, p. 32), a rede *multilayer* é uma estrutura paralela, constituída por um polinômio de segundo grau como função de ativação. O número de neurônios gerados, em cada camada escondida, depende do número de variáveis de entrada. Neste modelo, cada duas variáveis são combinadas para gerar um neurônio escondido, por exemplo:  $x_i$  e  $x_j$  são passados para a unidade  $k$  que pode ser construída por uma função de ativação do tipo:

$$y = \alpha_0 + \alpha_1 x_i + \alpha_2 x_j + \alpha_3 x_i x_j + \alpha_4 x_i^2 + \xi. \quad [9]$$

Os pesos são estimados usando-se um conjunto de treinamento e um dado valor de limiar, um critério de erro é usado para avaliar o ajuste destes neurônios gerados. Quando se tem  $n$  variáveis de entrada, a primeira camada gerará  $C_2^n$ , isto é,  $n(n-1)/2$  neurônios parciais. Desses, escolhem-se os que são superiores a um dado valor de limiar estabelecido, que passarão a compor a próxima camada. Este processo é repetido para sucessivas camadas até que o mínimo global para um dado critério de erro seja encontrado. Os critérios utilizados para finalizar o treinamento, citados por Valença (2005, p. 27), são: a camada de saída possui apenas um neurônio ou a camada adicionada não fornece ganhos significativos em relação à anterior.

### 3. MÉTODO

Para o desenvolvimento deste trabalho optou-se pela utilização de uma rede neural supervisionada, com algoritmo de seleção para frente. Dentre os modelos disponíveis, o método escolhido foi método GMDH, devido a sua característica híbrida de redes neurais e métodos estatísticos tradicionais, conforme cita Valença (2005, p. 24). A rede escolhida foi a

rede *multilayer* pela sua possibilidade de testar uma abrangente combinação de variáveis, bem como sua inferência estatística. O cálculo de cada neurônio será realizado através do método dos mínimos quadrados ordinários.

### 3.1. Conjunto de variáveis utilizadas no estudo

Kobayashi-Hillary (2007, p. 1), cita que BRIC ou BRIC's são termos utilizados para referir-se à combinação do Brasil, Rússia, Índia e China. O consenso geral é que o termo foi inicialmente utilizado em um relatório do banco de investimento Goldman Sachs, sendo tal sigla creditada à Jim O'Neill, economista do próprio banco.

O Produto Interno Bruto (PIB) nominal é o valor total final de bens e serviços produzidos por uma nação em um dado ano. De acordo com o Fundo Monetário Internacional (FMI, 2008), as cinco maiores economias do atual são: EUA (PIB de 13,840 trilhões de dólares), em segundo lugar encontra-se o Japão (PIB de 4,381 trilhões de dólares), em terceiro lugar a Alemanha (PIB de 3,320 trilhões de dólares), em quarto lugar a China (3,280 trilhões de dólares) e em quinto lugar o Reino Unido (PIB de 2,804 trilhões de dólares). Excluindo a China, os quatro países mais representativos na economia atual representam 44,60% do PIB mundial (54,584 trilhões de dólares). Os oito países analisados nesta pesquisa representavam 57,39% da riqueza gerada em 2007, o que justificou o interesse no estudo destes países.

Neste estudo foram utilizadas como variáveis o retorno mensal, na forma logarítmica, dos índices Ibovespa (Brasil), RTSI *Index* (Rússia), *Bombay Stock Exchange* (Índia), *Hang Seng Index* (China), S&P 500 (EUA), *Nikkei 225* (Japão), FTSE 100 (Reino Unido) e *Deutscher Aktien Index* (Alemanha), na forma da primeira diferença em  $t-1$ ,  $t-2$  e  $t-3$ , que formarão  $x_1$ ,  $x_2$  e  $x_3$ , respectivamente. A escolha dos índices deu-se pelo fato de serem considerados o mais importantes indicadores do desempenho médio das cotações do mercado de ações dos mercados em questão. Foram utilizados, na fase de treinamento, os retornos mensais de janeiro de 2000 até fevereiro de 2007, perfazendo 86 observações, a exceção do RTSI *Index* (Rússia) o qual foram utilizados 33 observações, referente aos retornos mensais de julho de 2004 até fevereiro de 2007. Na fase de teste foram simuladas 28 novas observações, ou 25% do tempo de treinamento, entre março de 2007 até junho de 2009.

Tsay (2002, p. 2), cita que a maioria dos estudos de séries temporais financeiras utiliza os retornos, em vez dos preços dos ativos em questão, neste sentido o autor comenta que existem duas razões principais para o uso dos retornos em estudos financeiros: em primeiro lugar, para a média dos investidores, o retorno dos ativos é uma medida adequada para a comparação entre oportunidade de investimentos e, em segundo lugar, as séries de retorno são mais fáceis de lidar do que uma série de preços, porque apresentam propriedades estatísticas mais atraentes. Dentre tais propriedades pode ser citada a ausência de tendenciosidade comum em séries de dados não estacionados. Tsay (2002, p. 11) comenta que a utilização de log-retornos em estudos financeiros é indicada pela hipótese de que os retornos dos ativos são independente e identicamente distribuídos (i.i.d.) com média  $\mu$  e variância  $\sigma^2$ .

Para encontrar o log-retorno inicia-se com a capitalização contínua, onde o preço de um título é dado por  $P_t = P_{t-1}e^r$  (uma vez que se está utilizando apenas um período, ou seja,  $t = 1$ ). Em que  $P_t$  é o preço da ação na data  $t$  e  $P_{t-1}$  é o preço da ação na data  $t-1$ .

A expressão pode ser reescrita da seguinte forma:

$$\frac{P_t}{P_{t-1}} = e^r. \quad [10]$$

Extraindo-se o logaritmo natural dos dois lados da igualdade obtêm-se:

$$r = \ln\left(\frac{p_t}{p_{t-1}}\right) \quad [11]$$

Inicialmente, às variáveis puras,  $x_1$ ,  $x_2$  e  $x_3$  serão acrescidas de uma constante. Na sequência, estas variáveis formarão a uma nova combinação multiplicando-se duas a duas, formando as novas variáveis:  $1, x_1, x_1^2, x_1x_2, x_1x_3, x_2x_3, x_2, x_2^2, x_3, x_3^2$ .

### 3.2. Cálculo da primeira camada da rede

Para iniciar o cálculo da primeira camada será utilizado o conceito do algoritmo combinatorial (COMBI) para a definição do número de neurônios a serem testados nesta camada. Haverá o desenvolvimento de combinações entre as dez variáveis, citadas anteriormente. Na equação [12], é apresentado o modelo de cálculo do total de combinações que serão testadas nesta camada:

$$N = C_{10}^1 + C_{10}^2 + C_{10}^3 + C_{10}^4 + C_{10}^5 + C_{10}^6 + C_{10}^7 + C_{10}^8 + C_{10}^9 + C_{10}^{10}, \quad [12]$$

onde N é o número de neurônios que formará a primeira camada da rede.

Posteriormente, haverá uma busca polinomial, semelhante ao método *stepwise*. A diferença básica está no critério de decisão que será adotado no modelo. O critério de decisão a ser adotado será o *Predict Squared Error (PSE)*, pois este além de levar em consideração o erro médio quadrático (MSE), leva em consideração a variância do erro e uma penalidade para os modelos mais complexos o que pode evitar o problema de *overfitting*. Na equação [13] é demonstrado o método de cálculo do PSE:

$$PSE = \left[ \frac{\sum_1^N (y_i - \hat{y}_i)^2}{n-2} \right] + 2\sigma_\varepsilon^2(p/n) \rightarrow \min, \quad [13]$$

Em que  $\sigma_\varepsilon^2$  é a variância do erro do modelo;  $p$  é o número de parâmetros no modelo;  $n$  o número de observações no arquivo de treinamento da rede;  $y_i$  são os valores tabelados da variável explicada e  $\hat{y}_i$  são os valores calculados de acordo com o modelo.

Serão selecionados os três melhores neurônios nesta camada segundo o critério *PSE*. Como o *PSE* apresenta uma relação direta entre o erro médio quadrático e a variância do erro, quanto menor o valor calculado em cada neurônio, maior será o ajustamento da variável prevista ( $\hat{y}_i$ ) com a variável observada ( $y_i$ ).

### 3.3. Cálculo nas camadas subsequentes

Os três neurônios selecionados formarão as novas variáveis de entrada para a próxima camada, isto é,  $w_1$ ,  $w_2$  e  $w_3$ . Haverá um novo desdobramento das variáveis e após haverá uma busca polinomial, entre as próximas combinações de variáveis, selecionando os três melhores neurônios a fim de diminuir o critério de decisão *Predict Squared Error (PSE)*. Esse processo de treinamento e seleção será então repetido para cada nova camada.

### 3.4. Critério para a finalização do treinamento e análise da previsão

Conforme destaca Valença (2005, p. 27), os critérios utilizados para finalizar o treinamento são: a camada adicionada não oferece ganhos significativos em relação à anterior ou o valor do *MSE* para o conjunto de treinamento é superior ao *MSE* da camada anterior, nesse caso o modelo ótimo estará na camada anterior. Para avaliar o sucesso das previsões,

conforme citam Ivakhnenko, Ivakhnenko e Müller (1993, p. 420), será utilizado a Equação [14]. Os resultados inferiores, ou iguais, à 0,5 serão considerados como adequados, os que estiverem entre  $0,5 < \delta^2 < 0,8$  serão considerados satisfatórios, os maiores que 1 serão considerados como desinformação sendo as modelagens consideradas como falhas, conforme citam Ivakhnenko e Ivakhnenko (1995, p. 532).

$$\delta_i^2 = \frac{\sum_1^N (y_i - \hat{y}_i)^2}{\sum_1^N (y_i - \bar{y})^2} \rightarrow \min, \quad [14]$$

Conforme citam Ivakhnenko e Ivakhnenko (1995, p. 532), com essa abordagem, cada modelagem a ser realizada em redes GMDH usará dois critérios. Em um primeiro momento, haverá uma exaustiva pesquisa entre todos os neurônios candidatos, conforme as combinações citadas no subitem 3.2.1, e um pequeno número de modelos cuja estrutura está perto de ótima são selecionados, neste caso será adotado o *PSE*, conforme é demonstrado na equação [13]. Então, apenas o neurônio selecionado como ótimo na última camada será testado em sua conformidade com um especial critério discriminatório, sendo, neste artigo, demonstrado na equação [14].

### 3.6. Análise dos erros e comparação entre mercados

A fim de se comparar a previsibilidade dos mercados, será utilizado o coeficiente de determinação amostral  $R^2$ , pois este procura medir a proporção ou percentual de variação de  $y$  previsto pelos modelos, conforme demonstrado na equação [15]:

$$R^2 = 1 - \frac{\sum_1^N (\hat{y}_i)^2}{\sum_1^N (y_i - \bar{y})^2}, \quad [15]$$

Outros dois indicadores serão utilizados: o erro quadrático médio (MSE) e o erro absoluto médio (MAE), que são demonstrados nas equações [16] e [17]:

$$MSE = \sqrt{\frac{1}{N} \sum_1^N (y_i - \hat{y}_i)^2}, \quad [16]$$

$$MAE = \frac{1}{N} \sum_1^N \left| \sqrt{y_i^2} - \sqrt{\hat{y}_i^2} \right|, \quad [17]$$

Adicionalmente será analisado o coeficiente de desigualdade de Theil, também chamado de  $U$ . O numerador de  $U$  é o *MSE*, mas a escala do denominador é tal que  $U$  estará sempre entre zero e um. Se  $U=0$  há um ajustamento perfeito da previsão com o valor observado. Se  $U=1$ , o desempenho da previsão do modelo é o pior possível. O coeficiente de desigualdade de Theil é demonstrado na equação [18].

$$U = \frac{\sqrt{\frac{1}{N} \sum_1^N (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{N} \sum_1^N (y_i)^2} + \sqrt{\frac{1}{N} \sum_1^N (\hat{y}_i)^2}}, \quad [18]$$

Além do coeficiente de desigualdade de Theil, é possível se determinar as proporções  $U^M$  e  $U^S$  que são denominadas respectivamente de proporção da tendenciosidade e proporção da variância que permitem decompor o erro em suas fontes características.



Conforme Pindyck e Rubinfeld (2004, p. 244), a proporção de tendenciosidade ( $U^M$ ) é uma indicação do erro sistemático, pois mede quanto os valores médios das séries simulada e efetiva se desviam uns dos outros. Qualquer que seja o valor do coeficiente de desigualdade ( $U$ ), espera-se que  $U^M$  seja próximo de zero. Um valor elevado de  $U^M$  (acima de 0,1 ou 0,2) seria preocupante porque indicaria a presença de tendenciosidade sistemática, de modo que seria necessário rever os modelos. Nas equações [19] e [20] são demonstradas a proporção de tendenciosidade e proporção de variância, respectivamente.

$$U^M = \frac{\left(\bar{y}^S - \bar{y}^A\right)^2}{(1/T)\sum\left(y_t^S - y_t^A\right)^2}, \quad [19]$$

$$U^S = \frac{\left(\sigma_S - \sigma_A\right)^2}{(1/T)\sum\left(y_t^S - y_t^A\right)^2}, \quad [20]$$

em que  $\bar{y}^S, \bar{y}^A, \sigma_S$  e  $\sigma_A$  são as médias e os desvios padrão dos valores previstos e observados, respectivamente.

A proporção de variância  $U^S$ , conforme citam Pindyck e Rubinfeld (2004, p. 244), indica a capacidade de replicar o grau de variabilidade na variável que interessa. Se  $U^S$  é alto, significa que a série efetiva flutuou muito enquanto a série simulada mostrou pouca flutuação, ou vice-versa. Isso também seria preocupante, e poderia levar à revisão dos modelos.

#### 4. RESULTADOS

Todos os cálculos foram realizados através do Excel e nenhum software estatístico foi utilizado. Um total de 1.109.955 neurônios foram testados, sendo 132.990 testados na previsão do Bovespa, 151.404 neurônios na previsão do índice Chinês, 122.760 neurônios na previsão do índice Indiano, 106.392 neurônios testados na previsão do índice Russo, 163.680 neurônios testados na previsão do Japonês Nikkei 225, 141.174 neurônios testados na previsão do S&P 500, 148.335 testados na previsão do FTSE 100, do Reino Unido e 143.220 neurônios testados na previsão dos 28 meses do *Deutscher Aktien Index*, da Alemanha.

##### 4.1. Resultados da fase de teste

Na Tabela 1, abaixo destacada, são demonstrados os resultados das previsões para  $t+1$  para o índice RTSI *Index* da Rússia.

**Tabela 1 – Resultados das previsões para o índice RTSI Index da Rússia.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
Camada 1	0,00001	-0,00239	0,42857	0,03939	0,14804	0,14267	0,00348	0,00079	1,66156
Camada 2	0,01450	-0,12043	0,42857	0,04505	0,15664	0,16049	0,00308	0,00059	1,90050
Camada 3	0,04212	0,20524	0,46429	0,03135	0,13599	0,11931	0,05074	0,00193	1,32251
Camada 4	0,01924	0,13871	0,50000	0,03465	0,13994	0,12706	0,02899	0,00112	1,46145
Camada 5	0,07085	0,26618	0,53571	0,02725	0,12206	0,10463	0,00486	0,00171	1,14947
Camada 6	0,07524	0,27429	0,53571	0,02828	0,12807	0,10415	0,00202	0,00100	1,19309
Camada 7	0,07527	0,27434	0,53571	0,02843	0,12772	0,10402	0,00106	0,00090	1,19926
Camada 8	0,07579	0,27531	0,53571	0,02828	0,12683	0,10379	0,00115	0,00095	1,19280
Camada 9	0,07407	0,27216	0,53571	0,02984	0,12967	0,10554	0,00011	0,00049	1,25877

Evidencia-se na Tabela 1 que, apesar dos demais indicadores terem melhorado ao longo das camadas, houve prejuízo na previsão deste índice, pois conforme o indicador de sucesso da previsão (demonstrado na equação [14]), todos os resultados foram considerados insatisfatórios.

Na Figura [2] é demonstrado o retorno do índice russo e sua respectiva previsão

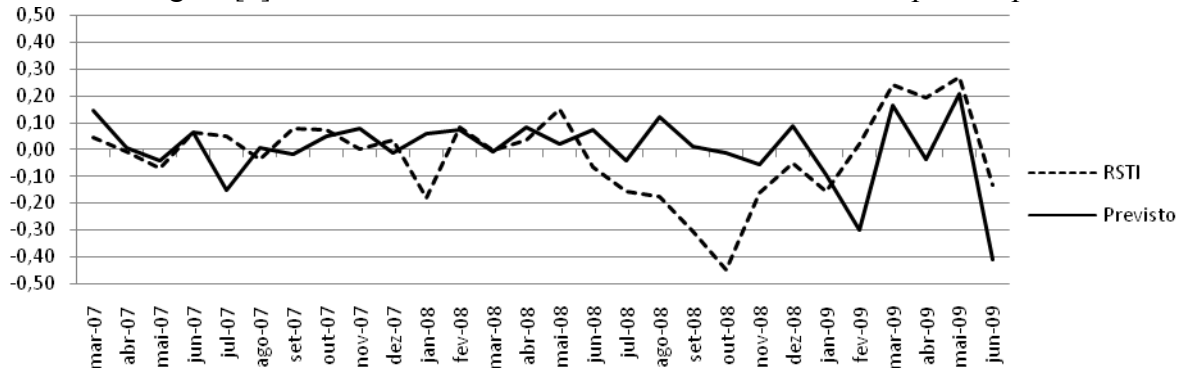


Figura 2 – Comparativo entre o retorno esperado e explicado para o RSTI *Index*.

**Tabela 2 – Resultados das previsões para o índice Hang Seng *Index*, da China.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
Camada 1	0,00004	0,00644	0,50000	0,01158	0,08899	0,06696	0,01377	0,00110	1,23438
Camada 2	0,01054	0,10267	0,57143	0,01106	0,08539	0,06411	0,00647	0,00109	1,17892
Camada 3	0,00044	-0,02102	0,53571	0,01483	0,09737	0,08471	0,00786	0,00096	1,58051
Camada 4	0,00032	-0,01801	0,57143	0,01554	0,10125	0,08975	0,00728	0,00106	1,65578
Camada 5	0,00036	0,01908	0,57143	0,01496	0,09893	0,08695	0,00687	0,00111	1,59418
Camada 6	0,00054	0,02316	0,57143	0,01487	0,09834	0,08569	0,01070	0,00105	1,58488
Camada 7	0,00593	0,07700	0,60714	0,01423	0,09667	0,07843	0,01201	0,00070	1,51705
Camada 8	0,00600	0,07743	0,60714	0,01423	0,09660	0,07840	0,01185	0,00070	1,51644
Camada 9	0,00598	0,07734	0,60714	0,01423	0,09653	0,07838	0,01244	0,00070	1,51640

De forma análoga à previsão do índice russo, a previsão índice Hang Seng *Index*, da China foi ineficiente, pois o indicador de Ivakhnenko apresentou resultados acima do que se poderia considerar como aceitável. O ponto significativo destes resultados é a evolução dos indicadores, demonstrando que houve um incremento na capacidade de previsão da rede à medida que foram sendo adicionadas camadas na modelagem.

Na Figura [3] é demonstrado o retorno do índice chinês e sua respectiva previsão.

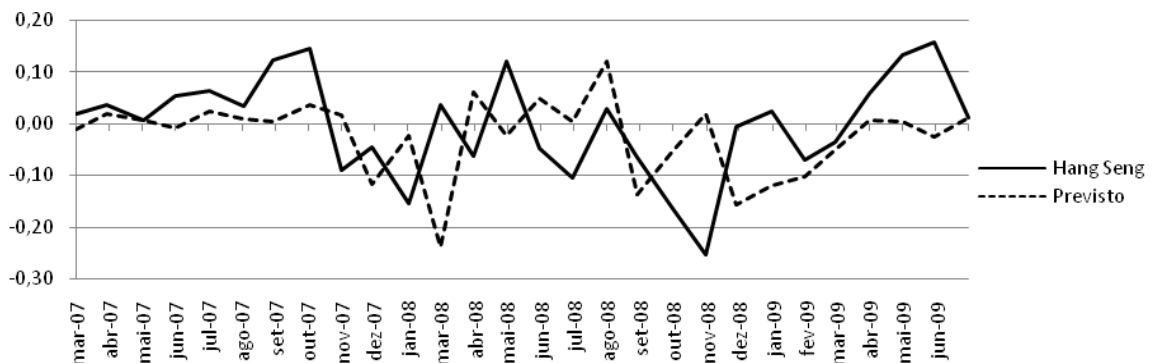


Figura 3 – Comparativo entre o retorno esperado e explicado para o índice Hang Seng.

Nas Figuras [2] e [3], percebe-se que a rede foi ineficaz para encontrar padrões na previsão dos índices russos (RSTI Index) e Chinês (Hang Seng Index), o que corrobora com o baixo valor para os índices  $R^2$  e do critério de Ivakhnenko (Equação [14]).

Na Tabela 3, são demonstradas as previsões do *Bombay Stock Exchange*, da Índia.

**Tabela 3 – Resultados das previsões para o índice *Bombay Stock Exchange*, da Índia.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
Camada 1	0,01360	-0,11662	0,57143	0,01744	0,10004	0,10082	0,01377	0,00110	1,46488
Camada 2	0,00013	0,01134	0,50000	0,01570	0,09868	0,09098	0,00647	0,00109	1,31848
Camada 3	0,01326	0,11515	0,60714	0,01454	0,09755	0,08306	0,00786	0,00096	1,22137
Camada 4	0,01221	0,11051	0,60714	0,01440	0,09672	0,08320	0,00728	0,00106	1,20964
Camada 5	0,02124	0,14576	0,64286	0,01382	0,09365	0,08032	0,00687	0,00111	1,16052
Camada 6	0,01576	0,12554	0,64286	0,01423	0,09536	0,08201	0,01070	0,00105	1,19550
Camada 7	0,00102	0,03201	0,60714	0,01654	0,10293	0,09113	0,01201	0,00070	1,38933
Camada 8	0,00080	0,02836	0,60714	0,01660	0,10299	0,09145	0,01185	0,00070	1,39402
Camada 9	0,00119	0,03449	0,60714	0,01651	0,10270	0,09092	0,01244	0,00070	1,38640

Percebe-se, na Tabela 3, que os resultados da previsão do índice indiano não foram significativos, conforme critério de Ivakhnenko, embora o critério de desigualdade de Theil, bem como a o indicador de tendenciosidade e de variância demonstram que as modelagens se mostraram não tendenciosas e relativamente ajustadas à previsão.

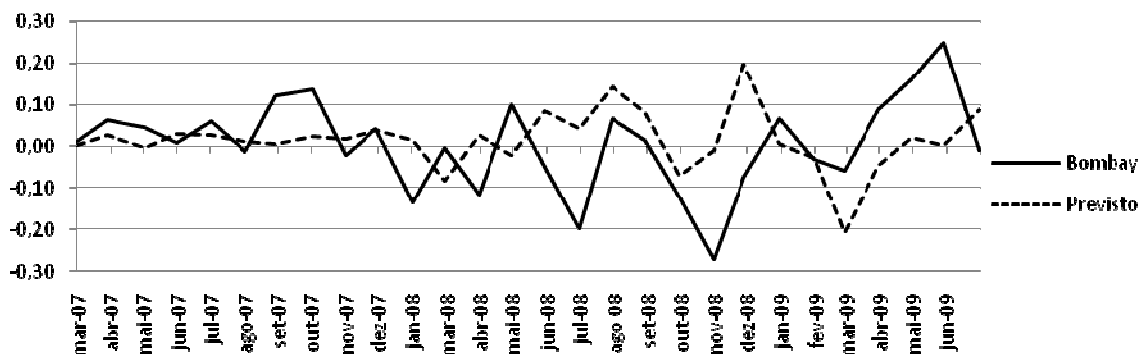


Figura 4 – Comparativo entre o retorno esperado e explicado para o índice *Bombay Stock Exchange*.

Abaixo são destacados os resultados para a previsão do índice IBOVESPA:

**Tabela 4 – Resultados das previsões para o índice Bovespa, do Brasil.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
Camada 1	0,23387	0,48360	0,67857	0,00655	0,06372	0,05763	0,02761	0,00237	0,83473
Camada 2	0,25718	0,50713	0,67857	0,00628	0,06408	0,05463	0,01530	0,00217	0,79962
Camada 3	0,27278	0,52228	0,67857	0,00618	0,06389	0,05348	0,01501	0,00210	0,78708
Camada 4	0,17357	0,41661	0,67857	0,00661	0,06656	0,05679	0,01096	0,00200	0,84191
Camada 5	0,21512	0,46381	0,67857	0,00632	0,06516	0,05329	0,00932	0,00178	0,80477
Camada 6	0,40904	0,63956	0,67857	0,00480	0,06087	0,03604	0,00182	0,00084	0,61138
Camada 7	0,41157	0,64153	0,67857	0,00478	0,06064	0,03588	0,00184	0,00084	0,60915
Camada 8	0,41069	0,64085	0,67857	0,00479	0,06075	0,03595	0,00170	0,00084	0,61001
Camada 9	0,41069	0,64085	0,67857	0,00479	0,06075	0,03595	0,00170	0,00084	0,61001

Na Tabela 4, é possível evidenciar que, no teste de Ivakhnenko, que todos os resultados foram considerados como satisfatórios, fato que pode ser evidenciado na evolução de todos os demais indicadores, em especial o  $R^2$  que evoluiu de 0,23 para a primeira camada para 0,41 para a última camada demonstrando a eficiência na previsão deste índice.

Abaixo é demonstrado o índice brasileiro e sua respectiva previsão.

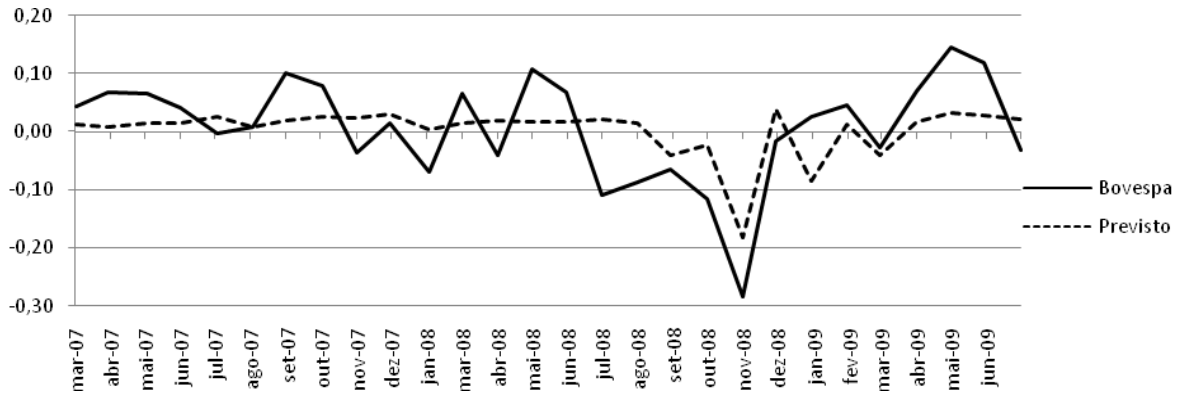


Figura 5 – Comparativo entre o retorno esperado e explicado para o índice BOVESPA.

**Tabela 5 – Resultados das previsões para o índice S&P 500, dos EUA.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
1º Camada	0,13199	-0,36330	0,32143	0,00692	0,06005	0,06668	0,02816	0,00015	1,83111
2º Camada	0,12509	-0,35368	0,39286	0,01742	0,07597	0,10566	0,06933	0,00030	4,60941
3º Camada	0,05258	-0,22929	0,35714	0,01776	0,08363	0,10546	0,00380	0,00053	4,69928
4º Camada	0,02646	-0,16267	0,39286	0,01324	0,07531	0,08585	0,04862	0,00017	3,50140
5º Camada	0,02611	-0,16159	0,42857	0,01323	0,07509	0,08583	0,04751	0,00017	3,50037
6º Camada	0,02582	-0,16070	0,42857	0,01322	0,07505	0,08576	0,04762	0,00017	3,49744
7º Camada	0,02580	-0,16063	0,42857	0,01322	0,07508	0,08577	0,04751	0,00017	3,49760
8º Camada	0,02580	-0,16063	0,42857	0,01322	0,07508	0,08577	0,04750	0,00017	3,49760
9º Camada	0,02580	-0,16063	0,42857	0,01322	0,07508	0,08577	0,04750	0,00017	3,49760

Analisando-se a Tabela [5], é possível perceber que a rede não encontrou padrão capaz de auxiliar na previsão do índice americano. Pelo contrário, na medida em que formam sendo incrementadas novas camadas na rede, o critério de informação de Ivakhnenko degenerou.

Na Figura [6] é demonstrado o índice americano e sua previsão.

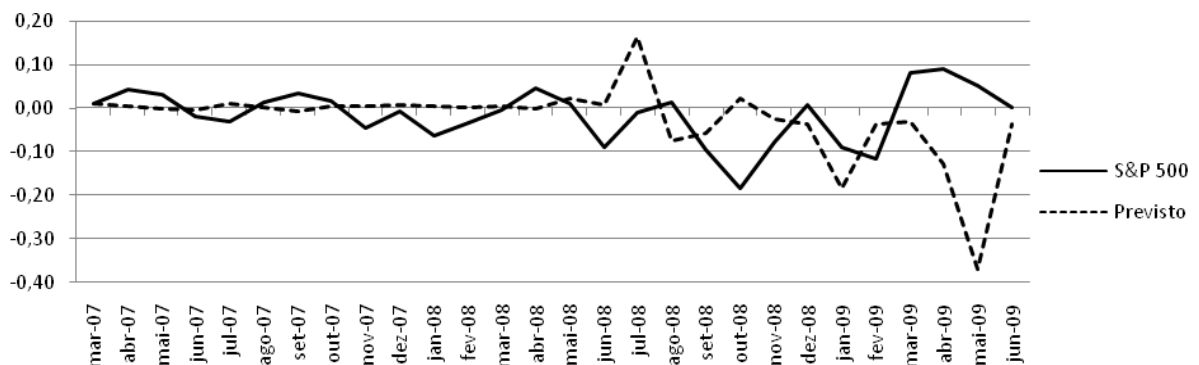


Figura 6 – Comparativo entre o retorno esperado e explicado para o índice S&P500.

Abaixo são destacados os resultados para a previsão do índice Nikkei 225, do Japão:

**Tabela 6 – Resultados das previsões para o índice Nikkei 225, do Japão**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
1° Camada	0,00935	-0,09668	0,50000	0,01067	0,07325	0,07520	0,01236	0,00009	1,79180
2° Camada	0,00635	-0,07968	0,46429	0,01011	0,07196	0,07270	0,01399	0,00012	1,69721
3° Camada	0,01951	-0,13968	0,53571	0,01008	0,07155	0,07549	0,00615	0,00020	1,69194
4° Camada	0,02209	-0,14863	0,50000	0,01025	0,07231	0,07618	0,00737	0,00018	1,71990
5° Camada	0,01971	-0,14041	0,46429	0,01012	0,07207	0,07552	0,00724	0,00019	1,69929
6° Camada	0,02077	-0,14413	0,50000	0,01018	0,07246	0,07579	0,00754	0,00019	1,70875
7° Camada	0,02153	-0,14674	0,50000	0,01020	0,07247	0,07594	0,00770	0,00018	1,71286
8° Camada	0,02019	-0,14208	0,50000	0,01016	0,07234	0,07560	0,00810	0,00018	1,70571
9° Camada	0,02030	-0,14248	0,50000	0,01016	0,07240	0,07562	0,00818	0,00018	1,70602

Ao se analisar a Tabela [6], percebe-se que, apesar da melhoria dos resultados pela inclusão de novas camadas, os resultados das previsões foram ineficientes, sendo os resultados considerados como desinformação pelo critério de Ivakhnenko (Equação [14]).

Na Figura [7] é demonstrado o índice japonês e sua previsão.



Figura 7 – Comparativo entre o retorno esperado e explicado para o índice Nikkei 225.

Na Tabela 7 são destacados os resultados para a previsão do índice *Deutscher Aktien Index*, da Alemanha.

**Tabela 7 – Resultados das previsões para o índice *Deutscher Aktien Index*, da Alemanha.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
1° Camada	0,27987	-0,52903	0,28571	0,00779	0,06850	0,07867	0,00551	0,00091	1,48115
2° Camada	0,08049	-0,28371	0,39286	0,00754	0,06794	0,07120	0,00930	0,00061	1,43338
3° Camada	0,06199	-0,24898	0,39286	0,00798	0,07027	0,07233	0,00225	0,00045	1,51788
4° Camada	0,05679	-0,23831	0,39286	0,00822	0,07221	0,07273	0,00228	0,00038	1,56323
5° Camada	0,03733	-0,19322	0,39286	0,00866	0,07445	0,07191	0,01102	0,00023	1,64699
6° Camada	0,06698	-0,25880	0,35714	0,00926	0,07613	0,07761	0,00033	0,00024	1,76047
7° Camada	0,03868	-0,19668	0,35714	0,01240	0,08202	0,08665	0,00468	0,00000	2,35727
8° Camada	0,05368	-0,23169	0,35714	0,01034	0,07861	0,08060	0,00172	0,00010	1,96584
9° Camada	0,05368	-0,23169	0,35714	0,01034	0,07861	0,08060	0,00172	0,00010	1,96584

Ao se analisar a Tabela [7] percebe-se, de forma análoga ao índice americano, que a inclusão de novas camadas foi ineficiente na previsão do índice alemão, sendo os resultados considerados como desinformação pelo critério de Ivakhnenko.

Na Figura [8] é demonstrado o índice alemão e sua previsão.

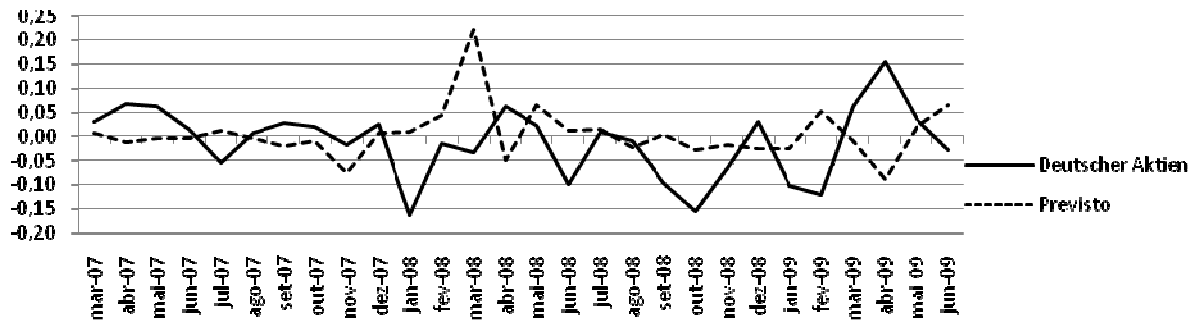


Figura 8 – Comparativo entre o retorno esperado e explicado para o índice *Deutscher Aktien Index*

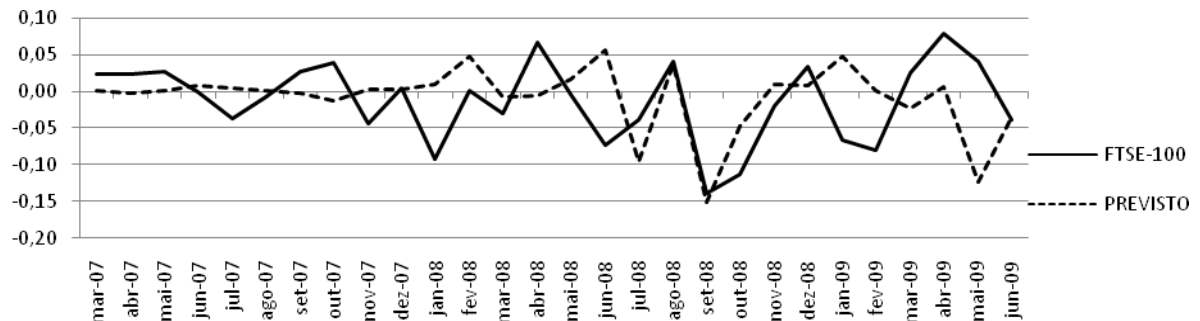
Na Tabela 8 são destacados os resultados para a previsão do índice FTSE 100, do Reino Unido.

**Tabela 8 – Resultados das previsões para o índice FTSE 100, do Reino Unido.**

Camada	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
1º Camada	0,11579	-0,34028	0,39286	0,00415	0,05105	0,05213	0,02729	0,00024	1,46432
2º Camada	0,01422	0,11925	0,57143	0,00404	0,04754	0,04087	0,05051	0,00003	1,42631
3º Camada	0,02354	0,15343	0,42857	0,00398	0,04711	0,03951	0,06800	0,00002	1,40476
4º Camada	0,05734	0,23947	0,50000	0,00370	0,04546	0,03671	0,02894	0,00002	1,30326
5º Camada	0,05136	0,22662	0,42857	0,00376	0,04679	0,03720	0,03629	0,00002	1,32755
6º Camada	0,05172	0,22742	0,42857	0,00386	0,04759	0,03803	0,01659	0,00001	1,36105
7º Camada	0,05113	0,22612	0,42857	0,00387	0,04764	0,03810	0,01664	0,00001	1,36431
8º Camada	0,05113	0,22612	0,42857	0,00387	0,04763	0,03810	0,01669	0,00001	1,36421
9º Camada	0,05113	0,22612	0,42857	0,00387	0,04763	0,03810	0,01669	0,00001	1,36421

É possível evidenciar, na Tabela 8, um incremento pouco significativo na capacidade de previsão do índice do Reino Unido. Os resultados foram considerados como desinformação pelo critério de Ivakhnenko (Equação [14]).

Na Figura [9] são demonstrados o índice FTSE 100 e sua previsão.



Na Tabela [9] são demonstrados os resultados das previsões dos países pesquisados:

**Tabela 9 – Resultados das previsões dos índices de cada um dos países analisados**

País	$R^2$	Correlação	Sinais	MSE	MAE	$U$	$U^M$	$U^S$	Ivakhnenko
------	-------	------------	--------	-----	-----	-----	-------	-------	------------

Brasil	0,41069	0,64085	0,67857	0,00479	0,06075	0,03595	0,00170	0,00084	0,61001
Rússia	0,07579	0,27531	0,53571	0,02828	0,12683	0,10379	0,00115	0,00095	1,19280
Índia	0,00119	0,03449	0,60714	0,01651	0,10270	0,09092	0,01244	0,00070	1,38640
China	0,00598	0,07734	0,60714	0,01423	0,09653	0,07838	0,01244	0,00070	1,51640
EUA	0,02580	-0,16063	0,42857	0,01322	0,07508	0,08577	0,04750	0,00017	3,49760
Japão	0,02030	-0,14248	0,50000	0,01016	0,07240	0,07562	0,00818	0,00018	1,70602
Alemanha	0,05368	-0,23169	0,35714	0,01034	0,07861	0,08060	0,00172	0,00010	1,96584
Reino Unido	0,05113	0,22612	0,42857	0,00387	0,04763	0,03810	0,01669	0,00001	1,36421

Na Tabela 9 destaca-se uma maior capacidade de previsão do índice brasileiro, quando comparado com os demais índices analisados nesta pesquisa. Neste sentido as modelagens GMDH foram capazes de encontrar não-aleatoriedade, no período analisado, para IBOVESPA, sendo ineficaz na previsão dos demais índices. A maior previsibilidade do índice brasileiro deu-se pela capacidade da rede em acertar o sinal para o período seguinte (decréscimo ou aumento do índice brasileiro).

## 5. CONSIDERAÇÕES FINAIS

As redes neurais, apresentam uma capacidade de encontrar padrões em dados ruidosos e imprecisos e aparentemente desconexos. Essa característica poderá ser utilizada em pesquisas financeiras, em especial de séries temporais. Neste sentido, a tentativa de utilização do método GMDH apresentou resultados significativos na previsão do índice IBOVESPA e não significativos para os demais índices. A previsibilidade do índice brasileiro pode ser explicada pelo estado de evolução do mercado brasileiro e pela qualidade do método de encontrar padrões não lineares para a previsão do índice.

A força desta modelagem pode ser expressa pela excelência na previsão em  $t+1$ , sendo ineficaz em modelagens para períodos posteriores a  $t+1$ . Cabe ressaltar ainda o fato de a modelagem ter sido feita toda em Excel o que se permite melhor desenvolver modelos de previsão, quando comparado a softwares que com códigos fonte fechados.

A maior capacidade de previsão do índice brasileiro (IBOVESPA) deu-se pelo fato do maior número de acerto de sinais (decréscimo ou aumento do índice brasileiro), quando comparado com os demais índices. A capacidade de previsão do sinal, em  $t+1$ , torna-se importante na medida em que se utiliza, entre outras estratégias de investimento, ações de curto prazo (compra e venda de títulos no curto prazo).

Como sugestão é possível indicar o uso de modelos com aprendizagem por correção de erros ou aprendizagem baseada em memória, ou ainda modelos de redes recorrentes já que esta modelagem foi feita através de redes alimentadas diretamente com múltiplas camadas. Para a criação de modelagens com aprendizagem por correção é possível indicar ainda a regra delta, ou regra de Widrow-Hoff. Para a modelagem com aprendizado baseada em memória é possível indicar o uso de conceitos da lógica difusa na determinação de padrões dos dados de entrada.

## REFERENCIAS BIBLIOGRÁFICAS

FERSON, W. in **Forecasting Expected Returns in the Financial Markets**. Ed. Elsevier, San Diego, Califórnia, 2007.

FMI. FUNDO MONETÁRIO INTERNACIONAL. Disponível em <http://www.imf.org>, Acesso 08/09/2009, 2008.

HAYKIN, Simon. **Redes Neurais: princípios e prática**. Segunda edição, Editora Bookman. Porto Alegre, 2001.

IVAKHNENKO, A. G. **Polynomial theory of complex systems**. IEEE Transaction on Systems, Man, Cybernetics, Vol. 1, pp 364-378, October, 1971.

IVAKHNENKO, A. G., IVAKHNENKO, G.A., e MULLER, J.A. **Self-Organization of Optimum Physical Clustering of the Data Sample for Weakened Description and Forecasting of Fuzzy Objects**. Pattern Recognition and Image Analysis, 1993, vol. 3, no. 4, pp. 415-422.

IVAKHNENKO, A. G., IVAKHNENKO. **The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH)**. Pattern Recognition and Image Analysis, 1995 Vol. 5, No. 4, 1995, pp.527-535

KOBAYASHI-HILLARY M. **Building a Future with BRICs: The Next Decade for Offshoring**, NY, New York City: Springer, 2007.

LUDWING Jr., O.; COSTA, E.M.M. **Redes Neurais: fundamentos e aplicações com programas em C**. Editora Ciência Moderna Ltda. Rio de Janeiro, 2007.

PALMA NETO, L.G.; NICOLETTI, M.C. **Introdução às Redes Neurais Construtivas**. Ed. UFSCar. São Carlos, 2005.



PINDYCK, R. S.; RUBINFELD, D. L. **Econometria: Modelos e previsões**. Editora Elsevier. Rio de Janeiro, 2004.

TOWELL, G.; SHAVLIK, J. **Knowledge-based artificial neural networks**. Artificial Intelligence, v.70, n.1-2., p.119-165, 1994.

TSAY, R. S. **Analysis of financial time series**. Ed. Wiley, New York, 2002.

VALENÇA, M. **Aplicando redes neurais: um guia completo**. Editora Livro Rápido, Olinda, 2005.