

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE GESTÃO E NEGÓCIOS  
BACHARELADO EM GESTÃO DA INFORMAÇÃO

LORENA FLÁVIA FERNANDES CUNHA

**CLASSIFICAÇÃO DE EMPRESAS EM GRUPOS DE RETORNO DAS AÇÕES  
UTILIZANDO REDES NEURAS ARTIFICIAIS**

UBERLÂNDIA – MG

LORENA FLÁVIA FERNANDES CUNHA

**CLASSIFICAÇÃO DE EMPRESAS EM GRUPOS DE RETORNO DAS AÇÕES  
UTILIZANDO REDES NEURAS ARTIFICIAIS**

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Gestão da Informação, da Universidade Federal de Uberlândia, como exigência parcial para a obtenção do título de Bacharel em Gestão da Informação.

Orientador: Prof. Dr. Vinícius Silva Pereira

Uberlândia - MG

2018

**CLASSIFICAÇÃO DE EMPRESAS EM GRUPOS DE RETORNO DAS AÇÕES  
UTILIZANDO REDES NEURAS ARTIFICIAIS**

Trabalho de Conclusão de Curso  
aprovado para a obtenção do título de Bacharel em  
Gestão da Informação da Universidade Federal de  
Uberlândia (MG), pela banca examinadora  
formada por:

---

**Prof. Dr. Vinícius Silva Pereira**  
Presidente da Banca - Orientador

---

**Prof. Dr. Antônio Sérgio Penedo**  
Membro

---

**Prof. Dr. Marcelo Fodra**  
Membro

Uberlândia – MG, 07 de dezembro de 2018

## AGRADECIMENTOS

Primeiramente agradeço à Deus por estar comigo em todos os momentos decisivos da minha vida.

Gostaria de agradecer a minha mãe, Vera Lúcia Fernandes, por ser uma guerreira, por ter batalhado todos esses anos para que eu conseguisse me formar. Uma batalha complicada, solitária... mas sempre estive forte com um sorriso no rosto. Minha melhor amiga, confidente, minha inspiração. Minha grande mãe.

Agradeço ao professor Vinícius pela orientação e aos demais professores da FAGEN.

Agradeço aos meus amigos Aleixa, Arthur Maurício, Gustavo Salomão e Rafael por estarem sempre comigo e não permitirem que a nossa amizade acabe, mesmo após a nossa formatura.

Agradeço aos meus grandes amigos Eduardo Marson e Leticia Pires por me apoiarem e nunca deixarem de me enviar forças, carinho e compreensão, mesmo de longe.

Agradeço a minha Tia Dalva, por ser a única parente paterna que acredita em mim e me deseja tudo que existe de melhor nesse mundo.

Enfim, agradeço a todos que fizeram com que sonho fosse realizado.

## RESUMO

A aplicação das redes neurais artificiais voltadas ao mercado financeiro tem sido utilizada em diversos estudos na área de Finanças e apresentado bons resultados com o objetivo de previsão e classificação. De outro lado, gestores financeiros e investidores utilizam-se dos números encontrados nas demonstrações financeiras para refletir a situação futura da empresa e tomar decisões. Partindo destas evidências este trabalho teve por objetivo classificar o retorno das ações de grupos de empresas de capital aberto brasileiras a partir de indicadores contábeis, financeiros e de mercado trimestrais de 2015 a 2017 processados em Redes Neurais. Como objetivos específicos foi proposto identificar as variáveis que influenciam na classificação do retorno das ações das organizações. Para tanto, coletou-se dados trimestrais de todas as empresas de capital aberto e de todos os indicadores disponíveis do Economatica entre 2015 e 2017. Na sequência aplicou-se uma análise fatorial nos indicadores para melhor agrupá-los em fatores e elaborou-se uma RNA otimizada, tendo como output o retorno trimestral das ações divididos em quartis e como input todos demais fatores. Os resultados encontrados demonstraram que a RNA utilizada no estudo conseguiu classificar corretamente 100% das empresas com retornos de primeiro e quarto quartis, mostrando a efetividade da técnica para segregar empresas de maior e menor retorno acionário a partir de indicadores.

**Palavras-Chaves:** Classificação; retorno das ações, redes neurais.

## **ABSTRACT**

The application of artificial neural networks focused on the financial market has been used in several studies in the area of Finance and presented good results for the purpose of forecasting and classification. On the other hand, financial managers and investors use the numbers found in the financial statements to reflect the future situation of the company and make decisions. Based on these evidences, the objective of this study was to classify the return of shares of Brazilian publicly traded companies based on quarterly accounting, financial and market indicators from 2015 to 2017 processed in Neural Networks. The specific objectives were to identify the variables that influence the return of the organizations' actions. A factorial analysis was then applied to indicators to better group them into factors and through SPSS software an optimized RNA was run, having the output of the quarterly return of shares divided into quartiles and the input of all other factors. The results showed that the RNA used in the study was able to properly classify 100% of the companies with returns of the first and fourth quartiles, showing the effectiveness of the technique to segregate companies with higher and lower shareholder returns from indicators.

**Keywords:** Prediction; profitability; neural networks.

## LISTA DE GRÁFICOS

<b>Gráfico 1:</b> Gráfico da Importância de variáveis .....	33
<b>Gráfico 2:</b> Gráfico da Importância de variáveis dos Fatores 2 e 10 .....	35
<b>Gráfico 3:</b> Gráfico da Importância utilizando variáveis fundamentalistas .....	38

## LISTA DE QUADROS

<b>Quadro 1:</b> Formas de eficiência de mercado.....	15
---	----

## LISTA DE TABELAS

<b>Tabela 1:</b> RMSE para diferentes configurações de rede neural.....	20
<b>Tabela 2:</b> Indicadores selecionados por especialistas.....	21
<b>Tabela 3:</b> Recursos de previsão média do Rede REF.....	21
<b>Tabela 4:</b> Recursos de previsão médios da rede do PCA.....	21
<b>Tabela 5:</b> Resultados da rede REF para o 39º e o 40º trimestres.....	22
<b>Tabela 6:</b> Resultados da rede PCA para o 39º e o 40º trimestres.....	22
<b>Tabela 7:</b> Resultados da rede REF para 2 empresas não treinadas.....	22
<b>Tabela 8:</b> Resultados da rede PCA para 2 empresas não treinadas.....	22
<b>Tabela 9:</b> Resultados para o Modelo de Previsão Padrão da Rede Neural: Modelo de Rácio.....	24
<b>Tabela 10:</b> Resultados para o Modelo de Previsão Padrão da Rede Neural: Rácio Financeiro e Modelo base-equity.....	24
<b>Tabela 11:</b> Variável Fundamentalista (Output RNA) .....	27
<b>Tabela 12:</b> Características do Setor Econômico da B3.....	28
<b>Tabela 13:</b> Características do Setor Econômico da B3 após normalização.....	28
<b>Tabela 14:</b> Classe das Ações das Empresas.....	28
<b>Tabela 15:</b> Variância Total Explicada.....	29
<b>Tabela 16:</b> Resumo do Processamento utilizando todos os fatores.....	31
<b>Tabela 17:</b> Informações de rede utilizando todos os fatores.....	32
<b>Tabela 18:</b> Resumo do modelo utilizando todos os fatores.....	32
<b>Tabela 19:</b> Classificação utilizando todos os fatores.....	33
<b>Tabela 20:</b> Resumo do processamento utilizando os fatores 2 e 10.....	34
<b>Tabela 21:</b> Resumo do modelo utilizando os fatores 2 e 10.....	34

<b>Tabela 22:</b> Classificação utilizando os fatores 2 e 10.....	34
<b>Tabela 23:</b> Variáveis fundamentalistas pertencentes aos fatores 2 e 10.....	35
<b>Tabela 24:</b> Resumo do processamento utilizando variáveis fundamentalistas.....	36
<b>Tabela 25:</b> Informações da rede composta pelas variáveis fundamentalistas.....	36
<b>Tabela 26:</b> Resumo do modelo utilizando variáveis fundamentalistas.....	37
<b>Tabela 27:</b> Classificação utilizando variáveis fundamentalistas.....	37

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>12</b>
<b>2. REFERENCIAL TEÓRICO.....</b>	<b>14</b>
<b>2.1. Hipótese da Eficiência do Mercado.....</b>	<b>14</b>
<b>2.2. Fianças Comportamentais.....</b>	<b>16</b>
<b>2.3. Estudos Empíricos.....</b>	<b>19</b>
<b>3. METODOLOGIA.....</b>	<b>25</b>
<b>4. ANÁLISE DOS RESULTADOS.....</b>	<b>27</b>
<b>5. CONSIDERAÇÕES FINAIS.....</b>	<b>38</b>
<b>6. REFERÊNCIAS.....</b>	<b>39</b>
<b>7. APÊNDICES.....</b>	<b>41</b>
<b>7.1. APÊNDICE A.....</b>	<b>41</b>

## 1. INTRODUÇÃO

A previsão econômica sempre foi um importante fator para a análise de mercado, reduzindo as incertezas na tomada de decisão por parte dos agentes envolvidos. Dessa forma, a análise financeira é importante ferramenta para a mensuração do desempenho econômico. Para isso, analistas e gestores de finanças utilizam-se dos números encontrados nas demonstrações financeiras para produzir índices, indicadores e previsões que venham a refletir a situação da empresa com relação à lucratividade, eficiência, crescimento e eficácia de suas ações, mantendo um maior controle sobre as possíveis decisões a serem tomadas, minimizando riscos de decisões erradas, que geram perdas futuras.

Soares e Galdi (2011), argumentam que a análise dos índices de desempenho do passado e do presente de uma empresa fornece a base para a previsão do desempenho futuro. Sendo assim, o propósito da análise desses índices é avaliar a eficácia das políticas da organização em diferentes áreas, possibilitando relacionar valores financeiros a fatores subjetivos e implícitos de gestão da empresa com maior nível de detalhe possível.

Baesso *et al.* (2008), afirmaram que “a redução da inflação e do risco no Brasil, a internacionalização dos mercados de capitais e a progressiva adesão do mercado brasileiro às regras contábeis internacionais estão entre os fatores que apontam para o aumento da eficiência do mercado de capitais brasileiro”.

Para Forti, Peixoto e Santiago (2009), uma vantagem comparativa para algum investidor só é possível caso o mesmo possua informações diferenciadas, que não estejam completamente incorporadas nos preços dos ativos. Além disso, é de suma importância ressaltar que quanto maior for a disponibilidade de informações ao mercado, maior será a tendência de que o mesmo seja eficiente, pois as informações novas ou exclusivas seriam cada vez mais dispendiosas para o investidor, eliminando, desta forma, possíveis ganhos adicionais desta informação.

A Hipótese da Eficiência de Mercado passou por revisões e ajustes decorrentes de várias críticas recebidas. Consequentemente, as finanças comportamentais possuem como objetivo o entendimento do processo através do qual os agentes tomam suas decisões, concluindo que os mesmos possuem tendência a utilizar regras simples para reduzir a complexidade das situações nas quais se envolvem. Com isso, as finanças comportamentais são consideradas contradição em relação a teoria dos mercados eficientes, para qual o preço atual reflete toda a informação relevante.

Nesse contexto, Haubert, Lima e Lima (2014) define finanças comportamentais como o estudo através do qual os investidores interpretam e agem frente às informações para decidir sobre investimentos. Segundo FAMÁ *et al.* (2008), essa linha de pesquisa visa identificar como as emoções e os erros cognitivos podem influenciar as decisões de investidores e como esses padrões de comportamento podem determinar mudanças no mercado.

O mercado de ações é um dos principais mecanismos para o desenvolvimento econômico, já que é um meio importante de captação de capitais. Deste modo, a tentativa de realizar a previsão do retorno das ações dessas organizações no mercado é de extrema importância, pois pode trazer grandes benefícios para os investidores, por aumentar o nível de informações sobre o mercado financeiro, minimizando a exposição ao risco.

De acordo com Borba e DILL (2007) a análise de índices de retorno das ações é uma das formas de avaliar a situação econômica da empresa, pois fornecem uma visão capaz de mensurar o estado atual da economia. Estes índices auxiliam no poder de decisão e planejamento do mercado financeiro, visto que suas oscilações podem causar significativos impactos ou rever algumas expectativas.

A tomada de decisão utilizando ferramentas matemáticas representa o avanço da tecnologia e do desenvolvimento interdisciplinar do conhecimento, pois casa estatística com gestão. Dessa forma, pode-se aplicar a técnica computacional denominada Redes Neurais Artificiais (RNAs).

Segundo Haykin (2001), as RNAs (Redes Neurais Artificiais) são modelos computacionais não lineares, que de forma similar ao cérebro humano, são organizadas através de elementos básicos de processamento denominados neurônios artificiais, que se interligam formando redes capazes de processar, armazenar e transmitir informações imputadas.

Dessa forma, Oliveira, Montini e Bergmann (2008), afirmam que as RNAs é uma ferramenta poderosa e flexível para realizar previsões de series temporais, quando uma seleção adequada de amostras de entradas e saídas, número de neurônios na camada intermediária, função de ativação e recursos computacionais estão disponíveis.

Sendo assim, o objetivo deste trabalho é classificar o retorno das ações de grupos de empresas de capital aberto brasileiras a partir de indicadores contábeis, financeiras e de

mercado trimestrais de 2015 a 2017 processados em Redes Neurais. Como objetivos específicos foi proposto identificar as variáveis que influenciam na classificação do retorno das ações das organizações.

## **2. REFERÊNCIAL TEÓRICO**

### **2.1. Hipótese da Eficiência de Mercado**

A Hipótese da Eficiência de Mercado (HEM) originou-se no período entre os séculos XIX e XX, quando foram desenvolvidos estudos sobre o comportamento dos preços de títulos. Em 1900, Louis Bachelier, formulou os primeiros relatos sobre a teoria do passeio aleatório, ou *Random Walk*, aplicada ao comportamento dos mercados, indicando que os preços dos títulos não possuíam ciclos regulares, ou seja, as variações dos preços são independentes uma das outras.

De acordo com Camargos & Barbosa (2003), em meados da década de 1960, a teoria foi formalizada matematicamente e incorporada em modelos econômicos. Nesta época, os trabalhos de Osborne (1959) e Roberts (1959), compararam as séries dos números aleatórios e dos preços correntes.

Pelo fato das discussões e teorias na época serem vagas, necessitava-se de uma teoria rigorosa. Assim, em 1970, Fama (1970) definiu a Hipótese da Eficiência de Mercado (HME), que passou a ser um dos pilares da Teoria de Finanças, movimentando até hoje grandes discussões a favor e contra a sua existência.

Para Jensen (1978), várias versões da Hipótese do Mercado Eficiente foram amplamente discutidas e testadas na literatura. As diferenças giram principalmente em torno da definição do conjunto de informações usado nesses testes.

Para Fama (1970), considerando um mercado no qual todas as informações estão disponíveis gratuitamente a todos os participantes do mercado, investidores, o preço atual de um título refletirá totalmente todas estas informações. Além disso, o retorno desses títulos, devem ter correlação serial igual a zero, ou seja, os investidores não podem utilizar essas informações como base de negociações lucrativas, já que a formação de preço obedecerá a um modelo aleatório. O quadro descreve as três formas de eficiência de mercado.

**Quadro 1:** Formas de eficiência de mercado

Forma de eficiência - Classificação antiga	Informação utilizada	Característica	Forma de eficiência - Classificação nova
<b>Fraca</b>	Retorno Passado	Não é possível obter retornos anormais, ou seja, acima da média do mercado, com base nas expectativas dos preços passados.	<b>Previsibilidade de Retornos Passados</b>
<b>Semi-forte</b>	Informações Públicas	Os preços se ajustam de forma eficiente a informações públicas (por exemplo, relatórios anuais, e etc.). Sendo assim, não é possível obter retornos anormais com base em tais informações.	<b>Estudos de Evento</b>
<b>Forte</b>	Informações Privadas	Da mesma forma que os casos citados anteriormente (retorno passado e informações públicas), os preços se ajustam as informações privadas.	<b>Testes de Informação Privada</b>

Fonte: Adaptado de Fama (1970).

Fama (1970) descreveu alguns testes necessários para identificar a eficiência dos mercados e, posteriormente, alterou o nome destes testes. Para cada forma, descrita no quadro, criou-se um teste, onde foram avaliados diversas vezes.

De acordo com Baesso *et al.* (2008), para identificar a forma fraca, os testes buscam indicar se os preços atuais das ações refletem as informações passadas. Este teste foi denominado de teste da previsibilidade dos retornos. Os testes da forma semiforte ou estudo de eventos, estão relacionados com a velocidade que os preços dos ativos se ajustam às novas informações públicas disponíveis, como as informações sobre lucros e dividendos das empresas. Por fim, os testes da forma forte ou teste de informações privadas, avaliam se os preços refletem todas as informações relevantes, não somente as públicas, para a formação de preços.

Larue e Sabuhoro (1997) afirmam que os preços devem refletir as informações pertinentes disponíveis e mudar com a chegada de novas informações. Além disso, um equilíbrio de preços caracterizado pelo uso racional da informação tem a vantagem da existência de erros sistemáticos. Sob essa ótica, a hipótese de eficiência de mercado pode ser considerada como uma hipótese conjunta de expectativas racionais e neutralidade de risco.

Para Forti, Peixoto e Santiago (2009), os custos de transação, a disponibilidade de informação e a liquidez dos títulos são fatores que podem influenciar a eficiência do mercado. Ao analisar essa afirmação, entende-se que a informação afeta a tendência, positivamente, de eficiência do mercado, pois as informações adicionais teriam custos, auxiliando, os investidores não adquirirem possíveis ganhos adicionais desta informação.

## 2.2. Finanças Comportamentais

A Teoria de Finanças Comportamentais surgiu, nesse cenário introduzindo estudos sobre o comportamento e irracionalidade do homem, aplicando, nestes estudos, conceitos relacionados com Psicologia e Sociologia à economia, para explicar as decisões financeiras dos indivíduos. Essa teoria visa identificar como as emoções e os erros cognitivos podem influenciar as decisões de investidores e, conseqüentemente, como essas mudanças podem afetar o comportamento do mercado (FAMÁ, R *et al.*, 2008).

Como mencionado anteriormente, a Teoria de Finanças Comportamentais relaciona conceitos de Psicologia e Sociologia à economia. Dessa forma, essa teoria analisa a combinação destes conceitos aplicados aos fatos da vida real, considerando que os agentes econômicos apresentam limitações racionais, que poderiam ser explicados por emoções e erros cognitivos (FILHO, 2011).

De acordo com Filho (2011), é de suma importância entender a diferença entre as Finanças Comportamentais e Finanças Neoclássicas. As Finanças Neoclássicas consideram o homem como um ser perfeitamente racional que possui todas as informações disponíveis e é capaz de analisar todas as informações e tomar uma decisão que as considere. Enquanto, as Finanças Comportamentais consideram que existem limites à racionalidade que ocorrem com frequência significativa para provocar resultados econômicos relevantes.

De acordo com Anache (2008), os limites à arbitragem e a psicologia são os dois pilares fundamentais sobre os quais se constrói a abordagem de Finanças Comportamentais. Tais pilares foram definidos por Shleifer e Summers (1990), definindo que as Finanças Comportamentais são comumente resumidas na aplicação da psicologia a finanças na tentativa de explicar as decisões financeiras dos indivíduos. Os pilares serão definidos a seguir.

- **Limites à arbitragem**

A hipótese de mercado eficiente, define que os preços dos ativos no mercado são equivalentes aos seus valores fundamentais, dado pela soma dos valores descontados dos seus fluxos de caixa esperado, ou seja, que os seus preços refletem seus valores corretos (ANACHE, 2008).

Para Yoshinaga *et al.* (2008), na visão de muitos economistas, ainda que parte dos agentes dos mercados financeiros atue de forma irracional, a Hipótese de Mercados Eficientes prevalecerá. Seu argumento baseia-se essencialmente em duas afirmações:

- Se houver um desvio dos preços dos ativos em relação a seus valores “corretos”, cria-se uma oportunidade atrativa de lucros; e
- Os agentes racionais aproveitarão rapidamente esta oportunidade, conduzindo os preços de volta a seus valores de equilíbrio, situação na qual o preço e o valor fundamental do ativo são iguais.

Yoshinaga *et al.* (2008) define a arbitragem como uma estratégia de investimento que oferece ganhos sem que haja risco ou custos adicionais. Os investidores racionais, ao agirem quando notam que os preços estão incorretos, são chamados de arbitadores, por buscarem um ganho sem risco. No entanto, Filho (2011) descreve que ao analisar o ponto de vista das Finanças Comportamentais, as estratégias adotadas pelos agentes racionais no mundo real não podem ser chamadas de arbitragem, pois, na maioria dos casos, são arriscadas.

Dentre os riscos a que os arbitadores estão expostos, destacam-se:

- Risco fundamental do ativo: Este risco está associado à dificuldade de se encontrar ativos que funcionem como substitutos perfeitos para os outros ativos objetos da arbitragem. Por exemplo, se o objeto da arbitragem é uma ação A, é preciso encontrar uma outra ação B cujos fluxos de caixa futuros sejam perfeitamente correlacionados com os da ação A. Todavia, em razão da dificuldade de encontrar ativos substitutos perfeitos, é impossível eliminar parte do risco fundamental do ativo (Shleifer e Vishny, 1995 apud Yoshinaga, 2008).
- O risco proveniente de movimentos dos investidores irracionais: Está relacionado com o risco de uma distorção de preços já detectada pelos arbitadores se torne ainda mais pronunciada no curto prazo. Ao existir uma distorção causada por comportamentos irracionais, é possível que esta aumente ainda mais em razão da continuidade dos referidos comportamentos. Neste caso, os arbitadores sofreriam severos prejuízos por um período de tempo que não se pode determinar com precisão, o que desestimularia a ação de arbitragem (YOSHINAGA *et al.*, 2008).

Por fim, é de suma importância destacar que o arbitrador terá, em muitos casos, que arcar com custos de implementação não triviais, diminuindo seu interesse pela operação de arbitragem.

- **Psicologia**

No segundo pilar da abordagem de Finanças Comportamentais, tem-se a psicologia. Após alguns estudos, diversos padrões de comportamento foram identificados. Como exemplos de heurística no processo de atualização de crenças dos indivíduos, encontram-se excesso de

confiança, otimismo, conservadorismo, perseverança, ancoragem com ajustamento insuficiente e disponibilidade. Em seguida, apresenta-se cada um deles em maior detalhamento.

Diversas pesquisas sobre o comportamento humano revelam que muitos indivíduos apresentam excesso de confiança nas próprias habilidades de estimação, não levando em consideração a verdadeira incerteza que existe no processo. As pessoas acreditam que suas estimativas em situações que envolvem incerteza são mais precisas do que realmente são (Yoshinaga, 2008).

Cabe ressaltar que o nível elevado de confiança nas atitudes leva o investidor a acreditar que possui inúmeras vantagens com a transação e durante a avaliação do bem em relação ao comportamento do mercado, além de manter a visão demasiada de ganhos futuros através das aquisições de ativos financeiros mal avaliados. Porém, em negociações comerciais, comumente, cada parte envolvida dispõe de informações distintas e muitas vezes contraditórias.

No entanto, reflexões e análises mais apuradas são postas em segundo plano, prevalecendo à confiança exagerada nas suas habilidades e informações disponíveis (Silva *et al.*, 2008).

Diversos padrões de comportamento foram identificados por diferentes pesquisadores, entre eles Weinstein (1980). O mesmo, em suas pesquisas, revelou que a maioria das pessoas, que participaram de sua pesquisa, fantasiavam suas habilidades e possibilidades, afirmando que poderiam fazer melhor do que estavam fazendo.

Lopes *et al.* (2011), descreve que o conservadorismo atua de maneira que os investidores apresentam lentidão para adquirir informações que contribuem rapidamente para a tomada de decisão, provocando a demora da formação dos preços, especialmente quando o que acontece no mercado difere das percepções dos investidores.

Em relação a perseverança, existe uma série de estudos que comprovam que as pessoas frequentemente têm opinião formada, o que provoca dois efeitos: elas relutam em procurar evidências que contrariem suas crenças e, ao encontrá-las, consideram-nas com excesso de ceticismo.

De acordo com Moura (2008), muitos buscam uma evidência confirmatória de um determinado evento, excluindo a busca por informações que não o confirmem. Isso acaba gerando um forte viés no processo decisório, pois não é possível saber se algo é verdadeiro sem haver uma checagem de sua possível não-confirmação.

Para FROMLET (2001), diversos outros elementos podem, também, influenciar a preferência dos investidores, por exemplo, risco/retorno, tamanho da empresa, valor de mercado/valor contábil. Dessa forma, modelos de precificação são ajustados de forma a

capturar o comportamento humano, que valorizam além das características utilitárias, mas indicadores de valor.

A ancoragem consiste na tendência de os indivíduos serem extremamente influenciados por crenças e opiniões prévias ou por eventos que ocorram com maior frequência (FILHO, 2011).

Para Yoshinaga (2004), essa tendência pode estar relacionada “em qualquer informação que lhes é fornecida, ajustando-o para obter uma resposta final”.

De acordo com Yoshinaga (2004), os vieses que emanam da heurística de disponibilidade estão relacionados com o fato de que os eventos mais frequentes são mais facilmente lembrados que os demais, e que os eventos mais prováveis são mais lembrados que os improváveis. Dessa forma, Lima Filho (2009), afirmou que o uso incorreto da heurística da disponibilidade pode levar a erros sistemáticos porque a disponibilidade de informações é afetada por outros fatores que não estão relacionados com a frequência do evento julgado.

Por fim, é de suma importância ressaltar que os estudos existentes comprovam que as heurísticas afetam o comportamento dos indivíduos de forma isolada, mas ainda não se conseguiu provar que estas anomalias afetam o mercado como um todo.

### **2.3. Estudos Empíricos**

As redes neurais artificiais, importante ramo da inteligência artificial, apresentam características muito relevantes para tarefas que envolvam previsão e classificação, tais como a capacidade de generalização, adaptação, correlação e aprendizado. Dessa forma, diversas áreas de conhecimento, como o agronegócio, biologia, economia e medicina, por exemplo, utilizam as RNAS em suas análises.

Fischmann *et al.* (2011), utilizaram as redes neurais artificiais como estratégia de previsão de preços no contexto do agronegócio, de forma a ser uma alternativa palpável aos gestores que, em diversas situações, precisam realizar negociações em mercados onde a incerteza está cada vez mais presente.

Neste trabalho, os autores citados anteriormente, implementaram o modelo na linguagem Java, utilizando os seguintes parâmetros do modelo: (1) critério de parada do algoritmo: 10.000 ciclos ou  $MSE < 0.005$ , o que ocorrer primeiro; (2) taxa de aprendizado = 0.8; (3) tamanho da amostra de treino: 140; (4) tamanho da amostra de validação: 20. A Tabela abaixo evidencia os valores estatísticos obtidos ao longo dos experimentos realizados com o modelo.

**Tabela 1.** RMSE para diferentes configurações de rede neural

Estadística	R <sup>2</sup>	MAPE	RMSE	Theil
Produto				
Soja	0,91097	4,37%	0,004906326	0,025717
Milho	0,772965	6,29%	0,007656769	0,040514
Boi Gordo	0,6923	5,73%	0,00732866380	0,033101
Trigo	0,87003	5,35%	0,0056250624	0,040849
Média	0,8115663	5%	0,00637900000	0,035045

Fonte: Fischmann *et al.* (2011).

Fischmann *et al.* (2011), salientam que os resultados obtidos podem ser considerados bastante relevantes, pois se demonstrou através das estatísticas R<sup>2</sup>, coeficiente de Theil (avalia o desempenho da previsão em relação à previsão ingênua ou trivial), RMSE (raiz do erro médio quadrático) e MAPE (erro percentual médio absoluto) que o modelo apresenta uma alta capacidade preditiva, podendo, portanto, afirmar-se que a prática do uso de redes neurais artificiais pode ser agregada ao processo decisório referente ao comportamento dos preços de *commodities* agrícolas e, conseqüentemente, a uma clarificação quantitativa a respeito dos riscos e incertezas que essa ação envolve.

Raposo e Cruz (200), discutiram a aplicação de uma combinação de técnicas de Redes Neurais e Lógica Fuzzy à análise fundamentalista do investimento em ações. Diversos pesquisadores usaram modelos de redes neurais de várias maneiras para prever previsões de estoque de curto e longo prazo, mas a maioria desses modelos usa indicadores técnicos como insumos. Nesta pesquisa, desenvolvida pelos autores, apresentam os resultados da aplicação de Redes Neurais Fuzzy para prever a evolução dos preços das ações de empresas brasileiras negociadas na Bolsa de Valores de São Paulo. A rede indica se um trader teria que manter, vender ou comprar uma ação usando uma combinação de informações extraídas de balanços (lançadas a cada três meses) e indicadores de mercado. A pesquisa foi realizada concentrada em um segmento da indústria brasileira, empresas do setor têxtil.

O resultado da pesquisa reduziu o conjunto de 52 indicadores disponíveis no banco de dados para nove, que foram usados por quase todas as fontes ou foram considerados muito importante pela alta econômica analistas entrevistados. A Tabela 2 listam esses indicadores e a classe a que pertencem.

**Tabela 2:** Indicadores selecionados por especialistas

<b>Fundamentalist Indicators</b>
Price/Profit
Firm Value/EBITDA
<b>Liquidity Indicators</b>
Current Liquidity Ratio
<b>Performance Ratios</b>
Asset Turnover
Operating Margin
Net Margin
Return on Assets
Return on Equity Basis
Return on Equity/Net Profit

Fonte: Raposo e Cruz (2004).

A Tabela 3 evidenciada abaixo, resume os recursos de previsão da rede obtida a partir de uma seleção de indicações de especialistas de mercado e referências, rotulada como REF.

**Tabela 3.** Recursos de previsão média do Rede REF.

<b>Prediction %</b>	<b>Total</b>	<b>Buy.</b>	<b>Sell</b>	<b>Keep</b>	<b>Not Class</b>
<b>Last trained trimester</b>	100	100	100	100	0
<b>One trimester ahead</b>	70.1	75.0	75.5	41.6	10.2
<b>Two trimesters ahead</b>	51.1	74.1	45.2	26.6	10.2

Fonte: Raposo e Cruz (2004).

A Tabela 4 mostra os resultados obtidos pela rede gerado a partir da análise de componentes principais e isso será referenciado como rede PCA. A última coluna (Não Classe) é a porcentagem de não insumos classificados. Estes números foram obtidos depois de testar as 20 empresas durante 5 trimestres.

**Tabela 4.** Recursos de previsão médios da rede do PCA.

<b>Prediction %</b>	<b>Total</b>	<b>Buy.</b>	<b>Sell</b>	<b>Keep</b>	<b>Not Class</b>
<b>Last trained trimester</b>	100	100	100	100	0
<b>One trimester ahead</b>	75.0	77.4	77.5	58.3	6.1
<b>Two trimestres ahead</b>	53.6	72.7	59.5	31.2	2.0

Fonte: Raposo e Cruz (2004).

Ao analisar os resultados, percebe-se que a rede PCA obteve os melhores resultados. Isso pode resultar do método que nós usado para agregar as escolhas dos especialistas. Outra contribuição para esses melhores resultados pode ser devido ao fato de que o método PCA leva em conta todos os indicadores econômicos ao criar a combinação linear. Tabelas 4 e 5 mostram que o os resultados para as metas de compra e venda eram razoáveis. Quanto ao alvo keep as redes não realizaram bem devido à baixa porcentagem de amostras deste tipo no início do

período considerado. Resultados parciais mostraram que quando apenas as duas últimas janelas de tempo foram consideradas, essa porcentagem é maior e o desempenho melhora. Tabela 5 e 6 mostra os resultados para os testes de 39 e 40 trimestres respectivamente. Outra conclusão é que a rede não tem informação suficiente para prever dois semestres à frente.

**Tabela 5.** Resultados da rede REF para o 39° e o 40° trimestres.

Prediction %	Total	Buy.	Sell	Keep	Not Class
39 <sup>TH</sup>	72.2	80.0	77.8	50.0	10.0
40 <sup>TH</sup>	84.2	100.0	81.3	100.0	5.0

Fonte: Raposo e Cruz (2004).

**Tabela 6.** Resultados da rede PCA para o 39° e o 40° trimestres.

Prediction %	Total	Buy.	Sell	Keep	Not Class
39 <sup>TH</sup>	75.0	75.0	80.0	66.7	15.0
40 <sup>TH</sup>	73.7	100.0	68.8	100.0	5.0

Fonte: Raposo e Cruz (2004).

As Tabelas 7 e 8 mostram os resultados dos testes aplicada a 2 empresas que não fazem parte do conjunto de empresas. Os resultados foram semelhantes aos obtidos de empresas treinadas indicando que tanto redes têm boa capacidade de generalização.

**Tabela 7.** Resultados da rede REF para 2 empresas não treinadas.

Prediction %	Total	Buy.	Sell	Keep	Not Class
One trimester ahead	70.0	80.0	66.6	50.0	0

Fonte: Raposo e Cruz (2004).

**Tabela 8.** Resultados da rede PCA para 2 empresas não treinadas.

Prediction %	Total	Buy.	Sell	Keep	Not Class
One trimester ahead	77.8	66.6	100.0	50.0	10.0

Fonte: Raposo e Cruz (2004).

Grigoryan (2015), utilizou as redes neurais para realizar a previsão direcional de preços, analisando os dados históricos. Em outras palavras, a análise se baseia na suposição de que variáveis de negociação passadas, como preço e volume, podem ajudar a prever tendências futuras do mercado. Dessa forma, o autor utilizou o conjunto de dados da bolsa de valores Nasdaq OMX Baltic para conduzir os experimentos. Todo o conjunto de dados pertencia ao período de 12 de março de 2012 a 30 de dezembro de 2014. Esse conjunto de dados incluiu o volume negociado, volume de negócios, abertura, fechamento, preços mais altos e mais baixos e 35 indicadores técnicos.

O modelo de classificação utilizado pelo autor é composto por dois estágios, onde foi realizado a combinação das técnicas de PCA e NARX. Primeiramente, a etapa de pré-processamento é aplicada aos dados que incluem: análise técnica para selecionar indicadores técnicos adequados, normalização de dados para ajustar e normalizar o conjunto de dados e análise de componentes principais para seleção de recursos e redução de dados. Após isso, o modelo NARX foi construído com base no subconjunto de recursos do PCA. Então, uma amostra foi treinada pela arquitetura paralela em série. Após a etapa de treinamento, a arquitetura paralela em série foi convertida em uma rede paralela, para executar a tarefa de previsão.

Grigoryan (2015), organizou os dados em dois conjuntos. Sendo que o primeiro conjunto de dados experimentais foi dividido em duas partes. A primeira parte (500 pares de observações) foi usado para as fases de formação, teste e validação. A segunda parte (200 pares de observações) foi reservada para a etapa de previsão. O segundo conjunto utilizou 550 amostras de dados para teste de treinamento e validação e 150 amostras de dados para tarefas de previsão. Como resultado, o método utilizado conseguiu prever valores próximos aos valores reais na maior parte do período de tempo. Os desempenhos de previsão são avaliados usando a medida de avaliação padrão chamada erro quadrático médio (MSE). O MSE da amostra com 150 dados resultou-se em 0.0011703 e a de 200 em 0.0034567. Dessa forma, os resultados experimentais mostraram que este método foi eficaz e eficiente para realizar a previsão dos preços das ações em comparação com outros estudos de pesquisa no campo da previsão do mercado de ações.

Atiya e Member (2001), descreve que a previsão de falências corporativas é um tópico importante, uma vez que pode ter um impacto significativo nas decisões e rentabilidade dos empréstimos bancários. Os Bancos precisam prever a possibilidade de inadimplência de uma contraparte potencial antes de conceder um empréstimo, tornando-se possível conduzir as decisões de empréstimo mais sólidas.

Para testar o sistema, foram coletados os dados históricos do padrão e de empresas americanas solventes. A cobertura de dados da empresa padronizada o período compreendido entre 1 mês e 36 meses antes da falência evento. As empresas foram selecionadas aleatoriamente (de entre todos solventes). Consideramos 716 empresas solventes e 195 firmas inadimplentes. As Tabelas abaixo evidenciem os resultados dos testes

**Tabela 9:** Resultados para o Modelo de Previsão Padrão da Rede Neural: Modelo de Rácio

Time to Default	# correct (in Sample)	# in Sample	% Correct (in Sample)	# correct (out of Sample)	# out of Sample	% Correct (out of Sample)
6 mnth or less	34	38	89.47	59	65	90.77
6 to 12 mnth	47	51	92.16	47	54	87.04
12 to 18 mnth	31	37	83.78	51	63	80.95
18 to 24 mnth	31	37	83.78	21	32	65.63
more than 24 mnth	14	25	56.00	24	42	57.14
<b>Total Defaulted</b>	<b>157</b>	<b>188</b>	<b>83.51</b>	<b>202</b>	<b>256</b>	<b>78.91</b>
<b>Solvent</b>	<b>258</b>	<b>303</b>	<b>85.15</b>	<b>343</b>	<b>413</b>	<b>83.05</b>
<b>Total</b>	<b>415</b>	<b>491</b>	<b>84.52</b>	<b>545</b>	<b>669</b>	<b>81.46</b>

Fonte: Atiya e Member (2001).

**Tabela 10:** Resultados para o Modelo de Previsão Padrão da Rede Neural: Rácio Financeiro e Modelo base-equity

Time to Default	# correct (in Sample)	# in Sample	% Correct (in Sample)	# correct (out of Sample)	# out of Sample	% Correct (out of Sample)
6 mnth or less	35	38	92.11	56	65	86.15
6 to 12 mnth	43	51	84.31	44	54	81.48
12 to 18 mnth	33	37	89.19	47	63	74.60
18 to 24 mnth	33	37	89.19	25	32	78.13
more than 24 mnth	19	25	76.00	28	42	66.67
<b>Total Defaulted</b>	<b>163</b>	<b>188</b>	<b>86.70</b>	<b>200</b>	<b>256</b>	<b>78.13</b>
<b>Solvent</b>	<b>276</b>	<b>303</b>	<b>91.09</b>	<b>372</b>	<b>413</b>	<b>90.07</b>
<b>Total</b>	<b>439</b>	<b>491</b>	<b>89.41</b>	<b>572</b>	<b>669</b>	<b>85.50</b>

Fonte: Atiya e Member (2001).

As Tabelas 9 e 10 mostram os resultados de ambos os sistemas, juntamente com uma quebra de acordo com o tempo até o padrão. Para o financeiro sistema de taxa obteve uma precisão de previsão de 84,52% para o conjunto na amostra e 81,46% para o conjunto fora da amostra.

Dessa forma, o próximo passo lógico para esta pesquisa é melhorar ainda mais o desempenho das RNAs para esta aplicação, talvez através de melhores métodos de treinamento, melhor seleção de arquitetura ou melhores entradas.

### 3. METODOLOGIA

A descrição formal das técnicas, procedimentos e métodos utilizados no desenvolvimento de uma pesquisa é conhecida como metodologia de pesquisa. Dessa forma, o presente estudo caracteriza-se como exploratório, pois estabelece critérios, métodos e técnicas para elaboração do mesmo, visando oferecer informações e orientar a formulação de hipóteses.

Os dados foram coletados por meio da Economatica®, selecionando os indicadores de mercado, financeiros e técnicos para as empresas de capital aberto para os anos 2015 a 2017,

dados trimestrais, formando assim uma amostra inicial de 564 companhias. Após a coleta foi realizado o empilhamento destes dados, totalizando 1440 observações.

As variáveis fundamentalistas (Anexo A) foram separadas em dois grupos, sendo que o Retorno do Fechamento foi utilizado como *target* e as demais variáveis como *input* da RNA. O *target*, Retorno do Fechamento, está descrito na Tabela 11.

**Tabela 11.** Variável Fundamentalista (Output RNA)

<b>Grupo</b>	<b>Indicador</b>	<b>Código</b>	<b>Fonte</b>
Indicadores Técnicos	Retorno	Retorno fechamento	Economática

Fonte: Dados da pesquisa.

Foi realizado a normalização, dividindo os valores pelo maior valor pertencente ao intervalo da variável. Após isso, os *missings* foram excluídos, pois é comum existirem, dentre as variáveis preditivas, algumas que possuem dados não preenchidos (*missings*), sendo necessário adotar algum procedimento para tratamento destas variáveis, por exemplo, a exclusão.

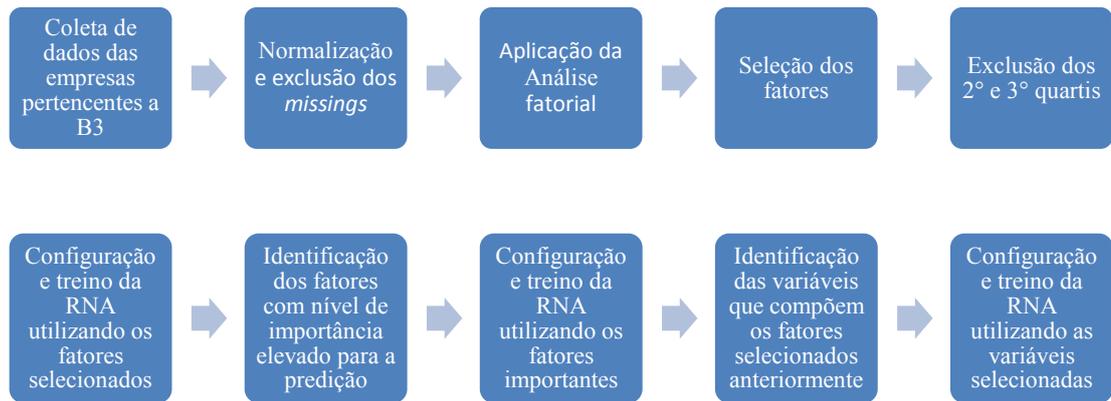
Após o procedimento de tratamento das variáveis, foram calculados os quartis, considerando o indicador técnico Retorno do Fechamento. Dessa forma, foi aplicado a Análise Fatorial para reduzir a grande quantidade de variáveis observadas em um número menor de fatores, pois nos testes realizados, as redes neurais não se comportaram de forma satisfatória.

Com os fatores selecionados, excluiu-se o segundo e quanto quartis e configurou-se a RNA com os fatores encontrados, utilizando como *input* os fatores e *output* a rentabilidade.

Sendo assim, testou-se a RNA e analisou-se quais fatores possuíam nível de importância elevado para a predição. Após identificar tais fatores, configurou-se novamente a RNA, utilizando como *inputs* os fatores selecionados e *output* a rentabilidade.

Após analisar a RNA configurada anteriormente, identificou-se quais variáveis compunham os fatores 2 e 10. Dessa forma, configurou-se a RNA utilizando como *input* tais variáveis identificadas e *output* a rentabilidade.

O processo abaixo descreve o passo a passo realizado neste presente trabalho.



Fonte: Elaborado pelo autor.

#### 4. ANÁLISE DOS RESULTADOS

Para a realização da análise dos resultados, foram coletados os indicadores financeiros, de mercado e técnicos das 564 empresas que compõem a carteira da B3, formando uma amostra com 77 variáveis. Quanto às características do setor econômico da B3, existe uma predominância de empresas do setor Financeiro como mostra a Tabela 12.

**Tabela 12.** Características do Setor Econômico da B3

	<b>Frequência</b>	<b>Porcentagem</b>
Financeiro e outros	131	23%
Consumo cíclico	117	21%
Utilidade pública	94	17%
Bens industriais	83	15%
Materiais básicos	61	11%
Consumo não cíclico	29	5%
Saúde	23	4%
Petróleo, gás e biocombustíveis	12	2%
Tecnologia da informação	7	1%
Telecomunicações	7	1%
<b>TOTAL</b>	<b>564</b>	<b>100%</b>

Fonte: Dados da pesquisa.

Conforme mencionado na metodologia, após a normalização dos dados, realizou-se a exclusão dos *missings*, formando uma nova amostra de dados com 120 companhias. Às Tabelas 13 e 14 demonstram às características do setor econômico da B3 da base de dados obtida.

**Tabela 13.** Características do Setor Econômico da B3 após normalização

	<b>Frequência</b>	<b>Porcentagem</b>
Consumo cíclico	33	28%
Utilidade pública	22	18%
Bens industriais	20	17%
Materiais básicos	15	13%
Consumo não cíclico	8	7%
Financeiro e outros	7	6%
Saúde	5	4%
Tecnologia da informação	4	3%
Petróleo, gás e biocombustíveis	3	3%
Telecomunicações	3	3%
<b>TOTAL</b>	<b>120</b>	<b>100%</b>

Fonte: Dados da pesquisa.

**Tabela 14.** Classe das Ações das Empresas

	<b>Frequência</b>	<b>Porcentagem</b>
ON	92	77%
PN	19	16%
UNT N2	2	2%
PNA	4	3%
PNB	3	3%
<b>TOTAL</b>	<b>120</b>	<b>100%</b>

Fonte: Dados da pesquisa.

Para Carvalho (2013), a análise fatorial é um método estatístico multivariado cujo objetivo é agrupar  $n$  variáveis aleatórias,  $X_1, \dots, X_n$ , em grupos formados por variáveis fortemente correlacionadas. Estes grupos constituem os fatores ou variáveis latentes. Os fatores são variáveis aleatórias não observáveis, preferencialmente em número inferior ao das variáveis originais. Utilizando o software SPSS (*Statistical Package for the Social Sciences*), programa usado para análises estatísticas, aplicou-se o método AF (Análise Fatorial) aos dados amostrais da presente pesquisa. A Tabela 15 evidencia o resultado da variação total explicada, sendo possível definir a quantidade de fatores dos dados.

**Tabela 15. Variância Total Explicada**

Autovalores iniciais			Somam de extração de carregamentos ao quadrado			
Componente	Total	% de variância	% cumulativa	Total	% de variância	% cumulativa
1	10,105	13,296	13,296	10,105	13,296	13,296
2	8,215	10,809	24,105	8,215	10,809	24,105
3	6,017	7,918	32,023	6,017	7,918	32,023
4	4,377	5,759	37,782	4,377	5,759	37,782
5	3,773	4,965	42,747	3,773	4,965	42,747
6	3,378	4,444	47,191	3,378	4,444	47,191
7	3,191	4,199	51,39	3,191	4,199	51,39
8	2,704	3,558	54,947	2,704	3,558	54,947
9	2,486	3,27	58,218	2,486	3,27	58,218
10	2,381	3,132	61,35	2,381	3,132	61,35
11	1,982	2,608	63,958	1,982	2,608	63,958
12	1,931	2,541	66,499	1,931	2,541	66,499
13	1,73	2,276	68,775	1,73	2,276	68,775
14	1,526	2,008	70,783	1,526	2,008	70,783
15	1,489	1,959	72,742	1,489	1,959	72,742
16	1,208	1,59	74,332	1,208	1,59	74,332
17	1,124	1,479	75,811	1,124	1,479	75,811
18	1,071	1,409	77,22	1,071	1,409	77,22
19	1,046	1,377	78,597	1,046	1,377	78,597
20	1,026	1,35	79,947	1,026	1,35	79,947
21	1,001	1,317	81,264	1,001	1,317	81,264
22	0,98	1,289	82,553			
23	0,913	1,202	83,754			
24	0,886	1,166	84,92			
25	0,877	1,153	86,073			
26	0,865	1,138	87,211			
27	0,817	1,075	88,286			
28	0,771	1,015	89,301			
29	0,708	0,932	90,233			
30	0,672	0,884	91,117			
31	0,642	0,845	91,962			
32	0,581	0,765	92,727			
33	0,545	0,717	93,444			
34	0,491	0,646	94,09			
35	0,417	0,549	94,639			
36	0,398	0,524	95,163			
37	0,35	0,461	95,624			
38	0,301	0,396	96,02			
39	0,289	0,38	96,4			
40	0,264	0,348	96,748			
41	0,252	0,332	97,08			
42	0,249	0,328	97,408			
43	0,222	0,292	97,7			
44	0,21	0,277	97,977			

45	0,189	0,248	98,225
46	0,177	0,233	98,458
47	0,159	0,209	98,667
48	0,142	0,186	98,854
49	0,118	0,155	99,009
50	0,105	0,138	99,147
51	0,091	0,12	99,267
52	0,086	0,113	99,38
53	0,059	0,077	99,457
54	0,054	0,071	99,528
55	0,046	0,06	99,588
56	0,045	0,059	99,647
57	0,037	0,049	99,696
58	0,033	0,044	99,74
59	0,03	0,039	99,779
60	0,025	0,033	99,812
61	0,021	0,028	99,84
62	0,02	0,027	99,867
63	0,017	0,023	99,889
64	0,016	0,021	99,911
65	0,015	0,02	99,931
66	0,014	0,019	99,949
67	0,014	0,019	99,968
68	0,012	0,016	99,984
69	0,005	0,007	99,991
70	0,002	0,003	99,994
71	0,001	0,002	99,996
72	0,001	0,002	99,998
73	0,001	0,001	99,999
74	0,001	0,001	100
75	0	0	100
76	1,35E-05	1,77E-05	100

Fonte: Dados da pesquisa.

Para definir a quantidade de fatores é necessário analisar as cargas fatoriais. A carga fatorial determina o fator que tem a maior influência em cada variável. As cargas fatoriais próximas de -1 ou 1 indicam que o fator influencia fortemente a variável. As cargas fatoriais próximas de 0 indicam que o fator tem uma influência fraca na variável. Dessa forma, pode-se afirmar que se tem 21 fatores, pois a partir deste os fatores não influenciam fortemente a variável.

Após aplicar a análise fatorial aos dados e definir os 21 fatores, realizou-se a exclusão do segundo e quarto quartil, pois tinha-se o intuito de considerar apenas os dados extremos, ou seja, os maiores e menores valores.

Construiu-se um modelo de rede neural para realizar a classificação do retorno das ações das empresas. De acordo com Corrar, Paulo e Filho (2009), na literatura não existe uma estrutura de RNA ótima para cada área de aplicação, ou seja, que apresente os melhores resultados. Dessa forma, para encontrar o melhor modelo, deve-se realizar tentativas empíricas.

Os dados utilizados no modelo construído foram separados em treinamento, teste e validação. Dessa forma, do total dos dados pertencentes a amostra, foram utilizados para treinamento 64,29%, para o teste 28,57% e validação 7,14%. No processamento da RNA, foram designados 472 casos da amostra para treinamento e 4 para teste. Os 244 casos da amostra foram excluídos, pois existem valores de variáveis de fator ou dependentes que não ocorrem na amostra de treinamento, conforme a Tabela 16.

**Tabela 16.** Resumo do Processamento utilizando todos os fatores

		<b>N</b>	<b>Porcentagem</b>
Amostra	Treinamento	472	99,20%
	Testes	4	0,80%
Válido		476	100,00%
Excluídos		244	
<b>Total</b>		<b>720</b>	

Fonte: Dados da pesquisa.

O SPSS seleciona qual a melhor função de ativação na camada oculta, o número de camadas ocultas, a função de ativação da camada de saída e a função de erro que consegue classificar os dados. Sendo assim, a RNA foi constituída com as informações listadas na Tabela 17.

**Tabela 17.** Informações de rede utilizando todos os fatores

Camada de entrada	Fatores		21 fatores
Camadas ocultas	Número de camadas ocultas		1
	Número de Unidades na Camada Oculta		14
	Função de ativação		Tangente hiperbólica
Camada de saída	Variáveis dependentes	1	Q_RF
	Número de unidades		2
	Função de ativação		Softmax
	Função de erro		Entropia cruzada

Fonte: Dados da pesquisa.

Realizando o treinamento da RNA, foi possível perceber que a rede acertou 100% da classificação no treinamento e nos testes, como pode ser evidenciado na Tabela 18.

**Tabela 18.** Resumo do modelo utilizando todos os fatores

Treinamento	Erro de entropia cruzada	0,191
	Porcentagem de predições incorretas	0,00%
	Regra de parada usada	1 passos consecutivos sem diminuição de erro
	Tempo de treinamento	00:12,7
Testes	Erro de entropia cruzada	3,68E-05
	Porcentagem de predições incorretas	0,00%

Fonte: Dados da pesquisa.

De acordo com Rocha (2015), no treinamento de uma RNA, encontra-se o erro entre as saídas calculadas e as saídas alvo desejadas. Dessa forma, o erro de entropia cruzada compara a distribuição de probabilidade gerada pelo modelo com o vetor da classificação real, ou seja, indica a distância entre o que a rede acredita que essa distribuição deve ser e o real. Sendo assim, pode-se perceber ao analisar a Tabela anterior que o erro foi baixo, visto que a RNA acertou a classificação.

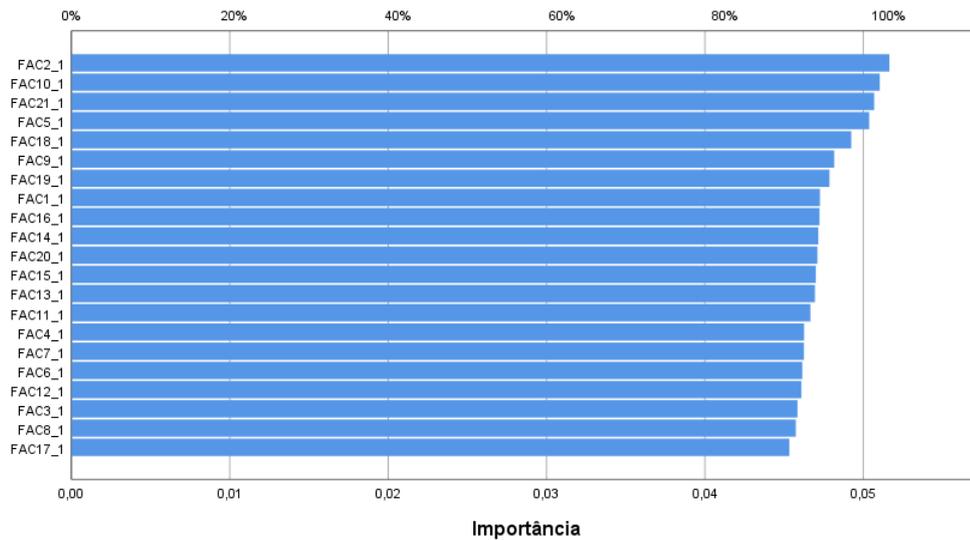
Ao analisar a Tabela 19, percebe-se que no treinamento a rede neural classificou 234 índices no primeiro quartil e 238 no quarto quartil, e na amostra de testes, a rede classificou corretamente em ambos quartis.

**Tabela 19.** Classificação utilizando todos os fatores

Amostra	Observado	1° QUARTIL	4° QUARTIL	Porcentagem Correta
Treinamento	1° QUARTIL	234	0	100,00%
	4° QUARTIL	0	238	100,00%
	<b>Porcentagem Global</b>	<b>49,60%</b>	<b>50,40%</b>	<b>100,00%</b>
Testes	1° QUARTIL	3	0	100,00%
	4° QUARTIL	0	1	100,00%
	<b>Porcentagem Global</b>	<b>75,00%</b>	<b>25,00%</b>	<b>100,00%</b>

Fonte: Dados da pesquisa.

O Gráfico 1 indica-nos a importância relativa e normalizada, ou seja, a capacidade explicativa de cada variável independente na classificação do retorno das ações das empresas. Podemos verificar que a variável com maior poder explicativo na classificação é o fator 2, com uma importância relativa de 100%, em seguida o fator 10, com uma importância relativa de 98,8%.

**Gráfico 1.** Gráfico da Importância de variáveis

Fonte: Elaborada pelo autor.

O Princípio da Parcimônia utiliza a premissa/hipótese estritamente necessária para a explicação de um fenômeno ou teoria. Com base neste princípio e com o intuito facilitar a análise do investidor e obter previsões que explique bem o relacionamento entre os dados, criou-se uma nova RNA constituída com os principais fatores, como variáveis independentes, e as mesmas configurações da rede anterior.

Ao processar os casos, 267 foram excluídos da análise, visto que os valores da amostra pertencentes ao teste e/ou validação tem valores de fator ou dependentes que não ocorreram na amostra de treinamento. Dessa forma, 99,10% da amostra foi utilizada no treinamento, 0,70% nos testes e 0,20% na validação, conforme a Tabela 20.

**Tabela 20.** Resumo do processamento utilizando os fatores 2 e 10

		<b>N</b>	<b>Porcentagem</b>
Amostra	Treinamento	449	99,10%
	Testes	3	0,70%
	Validação	1	0,20%
Válido		453	100,00%
Excluídos		267	
<b>Total</b>		<b>720</b>	

Fonte: Dados da pesquisa.

Na Tabela 21 encontram-se os resultados dos erros de entropia cruzada do treinamento e do teste, onde constatamos que apesar de haver erros, a RNA não teve previsões incorretas.

**Tabela 21.** Resumo do modelo utilizando os fatores 2 e 10

<b>Treinamento</b>	Erro de entropia cruzada	0,31
	Porcentagem de predições incorretas	0,00%
	Regra de parada usada	1 passos consecutivos sem diminuição de erros
	Tempo de treinamento	00:00,8
<b>Testes</b>	Erro de entropia cruzada	3,42E-06
	Porcentagem de predições incorretas	0,00%
<b>Validação</b>	Porcentagem de predições incorretas	0,00%

Fonte: Dados da pesquisa.

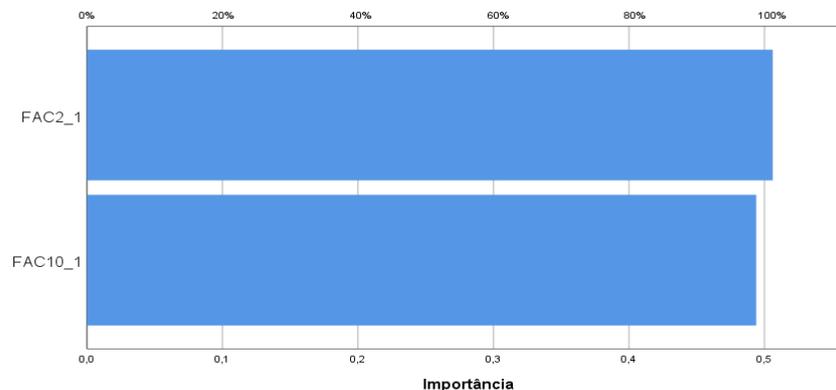
Analisando a Tabela 22 é possível observar a classificação correta por quartil das amostras de treinamento, teste e validação.

**Tabela 22.** Classificação utilizando os fatores 2 e 10

Amostra	Observado	1° QUARTIL	4° QUARTIL	Porcentagem Correta
Treinamento	1° QUARTIL	222	0	100,00%
	4° QUARTIL	0	227	100,00%
	<b>Porcentagem Global</b>	<b>49,40%</b>	<b>50,60%</b>	<b>100,00%</b>
Testes	1° QUARTIL	1	0	100,00%
	4° QUARTIL	0	2	100,00%
	<b>Porcentagem Global</b>	<b>33,30%</b>	<b>66,70%</b>	<b>100,00%</b>
Validação	1° QUARTIL	1	0	100,00%
	4° QUARTIL	0	0	0,00%
	<b>Porcentagem Global</b>	<b>100,00%</b>	<b>0,00%</b>	<b>100,00%</b>

Fonte: Dados da pesquisa.

O Gráfico 2 demonstra a importância normalizada das variáveis. Dessa forma, percebe-se que o grau de importância das variáveis foi listado na mesma ordem que o Gráfico 1.

**Gráfico 2.** Gráfico da Importância de variáveis dos Fatores 2 e 10

Fonte: Elaborada pelo autor.

Sendo assim, sabe-se que para verificar a classificação do retorno das ações da empresa de modo mais eficaz, deve-se analisar os fatores 2 e 10, pois utilizando estes, foi possível classificar 100% das empresas.

A partir da análise fatorial identificou-se as variáveis fundamentalistas que compunham os fatores 2 e 10. Para tal identificação, selecionou-se os maiores e menores valores de cada variável e realizou-se a comparação com os respectivos fatores. A Tabela 23 evidencia as variáveis selecionadas.

**Tabela 23.** Variáveis fundamentalistas pertencentes aos fatores 2 e 10

VARIÁVEL	MAIOR	MENOR	FATOR 2	FATOR 10
ExgAt	0,731		0,731	
DLPL	0,704		0,704	
Est_Cap_PL	0,698		0,698	
Div_Tt_Bruta	0,692		0,692	
DivBrPatr	0,69		0,69	
ExigPL	0,688		0,688	
Div_Tt_Lq	0,657		0,657	
AFxPL	0,554		0,554	
Cap_Empl	0,542		0,542	
DivBrAtivo	0,54		0,54	
GiroPL	0,537		0,537	
InvestCap	0,522		0,522	
EstCapVMerc	0,462		0,462	
Capex	0,455		0,455	
Depr_e_Amor	0,45		0,45	
Pmform	0,389		0,389	
EBITDivBr		-0,198	-0,198	
RenPatin		-0,227	-0,227	
RenPatmed		-0,268	-0,268	
RenPatfim		-0,287	-0,287	
RentAt		-0,351	-0,351	
LiqSec	0,589			0,589
LiqCor	0,648			0,648
LiqGer	0,569			0,569

Fonte: Dados da pesquisa.

Com as variáveis selecionadas, aplicou-se o modelo de RNA utilizado na presente pesquisa. Foi utilizado 93,6% da amostra para treinamento, 4,80% para testes e 1,60% para a validação. Conforme dito anteriormente, alguns casos são excluídos por não ocorrerem na amostra de treinamento. Sendo assim, conforme a Tabela 24, foram excluídos 221 casos.

**Tabela 24.** Resumo do processamento utilizando variáveis fundamentalistas

		<b>N</b>	<b>Porcentagem</b>
Amostra	Treinamento	467	93,60%
	Testes	24	4,80%
	Validação	8	1,60%
Válido		499	100,00%
Excluídos		221	
<b>Total</b>		<b>720</b>	

Fonte: Dados da pesquisa.

A Tabela 25 é composta pelas informações da RNA. Ao analisa-la é possível perceber que a rede foi constituída com a mesma estrutura das RNAs testadas anteriormente.

**Tabela 25.** Informações da rede composta pelas variáveis fundamentalistas

Camada de entrada	Fatores	20	Variáveis fundamentalistas
Camadas ocultas	Número de camadas ocultas		1
	Número de Unidades na Camada Oculta 1a		21
Camada de saída	Função de ativação		Tangente hiperbólica
	Variáveis dependentes	1	Q_RF
	Número de unidades		2
	Função de ativação		Softmax
	Função de erro		Entropia cruzada

Fonte Dados da pesquisa.

O resumo do modelo, evidenciado na Tabela 26, demonstra que houve 0,40% de predições incorretas no treinamento e 0 % no teste e validação respectivamente.

**Tabela 26.** Resumo do modelo utilizando variáveis fundamentalistas

<b>Treinamento</b>	Erro de entropia cruzada	2,776
	Porcentagem de predições incorretas	0,40%
	Regra de parada usada	Mudança relativa no critério de erro de treinamento (,0001) obtida
	Tempo de treinamento	00:35,0
<b>Testes</b>	Erro de entropia cruzada	0
	Porcentagem de predições incorretas	0,00%
<b>Validação</b>	Porcentagem de predições incorretas	0,00%

Fonte: Dados da pesquisa.

A porcentagem de predições incorretas no treinamento listada na Tabela anterior, pode ser confirmada ao analisar a Tabela 27, pois no primeiro e quarto quartil houve uma

classificação correta no treinamento de 99,60% e nas demais amostras foram classificadas corretamente.

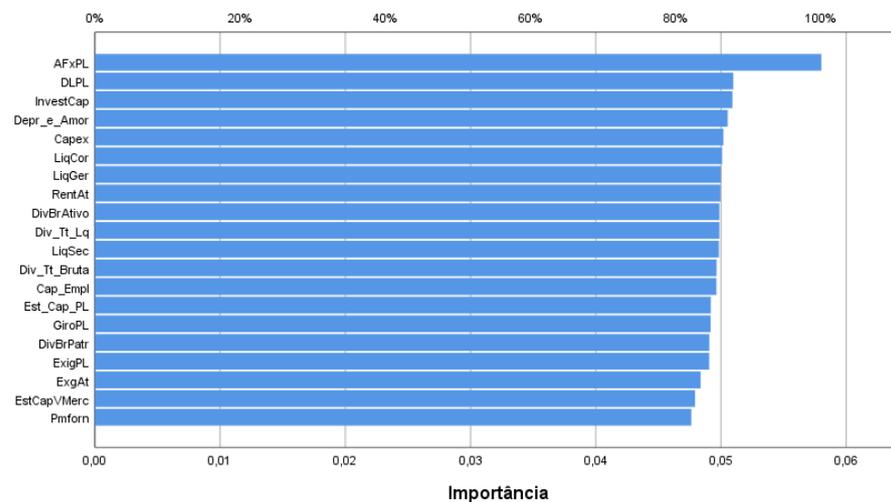
**Tabela 27.** Classificação utilizando variáveis fundamentalistas

Amostra	Observado	1° QUARTIL	4° QUARTIL	Porcentagem Correta
Treinamento	1° QUARTIL	240	1	99,60%
	4° QUARTIL	1	225	99,60%
	<b>Porcentagem Global</b>	<b>51,60%</b>	<b>48,40%</b>	<b>99,60%</b>
Testes	1° QUARTIL	18	0	100,00%
	4° QUARTIL	0	6	100,00%
	<b>Porcentagem Global</b>	<b>75,00%</b>	<b>25,00%</b>	<b>100,00%</b>
Validação	1° QUARTIL	5	0	100,00%
	4° QUARTIL	0	3	100,00%
	<b>Porcentagem Global</b>	<b>62,50%</b>	<b>37,50%</b>	<b>100,00%</b>

Fonte: Dados da pesquisa.

Como mencionado anteriormente, o gráfico que demonstra a importância relativa e normalizada das variáveis, revela a capacidade de cada variável independente prever o retorno das ações das empresas. Dessa forma, de acordo com o Gráfico 3, percebe-se que a variável AFxPL (Ativo Fixo/ Pat Liq) possui 100% de importância, seguido da variável DLPL (Dívida Líquida/ Patrimônio Líquido) com 87,9%.

**Gráfico 3.** Gráfico da Importância utilizando variáveis fundamentalistas



Fonte: Elaborada pelo autor.

Os resultados da classificação de retorno das ações encontrados na relação observados versus preditos, podem ser considerados efetivos, visto que em grande parte dos momentos os

dados preditos se aproximam dos observados e em outros momentos partilham os mesmos resultados.

## **5. CONSIDERAÇÕES FINAIS**

O presente estudo visou analisar as redes neurais com o objetivo de prever, através de modelos baseados em Redes Neurais Artificiais, o comportamento dos índices de retorno das ações das empresas de capital aberto.

Este estudo ilustra a aplicação prática de técnicas de classificação de padrões que podem ser utilizadas para desenvolver e aprimorar qualquer tipo de ferramenta de apoio à escolha de investimentos, ou seja, qualquer processo que envolva tomada de decisão.

Os resultados encontrados demonstraram que a RNA utilizada no estudo conseguiu classificar corretamente 100% das empresas com retornos de primeiro e quarto quartis, mostrando a efetividade da técnica para segregar empresas de maior e menor retorno acionário a partir de indicadores.

No que tange ao ponto de vista acadêmico e devido à grande aplicabilidade de inteligência computacional a problemas ligados ao mercado financeiro, este estudo pode ser utilizado como base para estudos mais aprofundados.

## 6. REFERÊNCIAS

- ANACHE, M, C, A. **Finanças Comportamentais** - uma avaliação crítica da moderna teoria de finanças. Dissertação (Dissertação em Economia) - UFES. Vitória. 2008.
- ATIYA, A, F; MEMBER, S. Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results. **Ieee Transactions On Neural Networks**, v. 12, n. 4, p.929-935, 2001.
- BAESSO ET AL. **Teste da Hipótese de Eficiência do Mercado no Brasil**: uma Aplicação de Filtros Ótimos. XXXII Encontro da ANPAD, Rio de Janeiro, 2008.
- BORBA, J.A.; DILL, R. Análise da Retorno das ações de Empresas: uma Abordagem Baseada na Lógica Nebulosa (Fuzzy Logic). **RAC-Eletrônica**. v. 1, n. 1, art. 4, p. 47-66, Jan./Abr. 2007.
- CAMARGOS, M, A; BARBOSA, F, V. Teoria e evidência da eficiência informacional do mercado de capitais brasileiro. **Caderno de Pesquisa em Administração da FEA/USP**, v. 10, n.1, p. 41-55, São Paulo, 2003.
- FAMA, E, F. Efficient capital markets: a review of theory and empirical work. **The Journal of Finance**, v. 25, n. 2, p. 383-417, Cambridge, May, 1970.
- FAMÁ, R et.al. Hipótese de mercados eficientes e finanças comportamentais: as discussões persistem. **Facef pesquisa desenvolvimento e gestão**, v.11, n.1, p. 5-17, São Paulo, 2008.
- FISCHMANN et al. Utilização de redes neurais artificiais como estratégia de previsão de preços no contexto de agronegócio. **Revista de Administração e Inovação**, v.8, n.4, p. 6-26, 2011.
- FORTI, C, A, B; PEIXOTO, F, M; SANTIAGO, W, P. Hipótese da eficiência de mercado: um estudo exploratório no mercado de capitais brasileiro. **Gestão & Regionalidade**, v.25, n.75, p. 45-56, São Paulo, 2009.
- FROMLET,H. Behavioral finance and Practical Application. **Business Finances**, V.36, n.3, July 2001.
- GRIGORYAN, H. Stock Market Prediction Using Artificial Neural Networks. Case Study of TALIT, Nasdaq OMX Baltic Stock. **Database Systems Journal**, v.6, n.2, p. 14-23, 2015.
- HAUBERT, F,L, C; LIMA, C, R, M E LIMA, M, V, A. Finanças comportamentais: uma investigação com base na teoria do prospecto e no perfil do investidor de estudantes de cursos

stricto sensu portugueses. **Revista de Ciências da Administração**, v. 16, n. 38, p. 183-195, 2014.

HAYKIN, S. **Redes Neurais: Princípios e Práticas**. (2ed.). Bookman.

JENSEN, M. C. Some Anomalous Evidence Regarding Market Efficiency. **Journal of Finance**, v. 6, p. 95-101, 1978.

LIMA FILHO, F; R, N. **Quanto mais faço, mais erro? Um estudo sobre associação entre prática de controladoria, cognição e heurísticas**. Dissertação de Mestrado. Universidade Federal da Bahia. Salvador, 2009.

LOPES et.al. **Finanças Pessoais: como Administrar Consumo e Gerar Poupança**. VIII Simpósio de Excelência em Gestão e Tecnologia. 2011.

MOURA, B, M. **Um indicador de desempenho para seleção de ativos das empresas de celulose e papel no mercado Financeiro**. Dissertação de Mestrado em Ciências Florestais, Departamento de Engenharia Florestal, Universidade de Brasília Brasília-DF, 2008.

OLIVEIRA, M, A; MONTINI, A, A; BERGMANN, D, R. Previsão de retornos de ações dos setores financeiro, de alimentos, industrial e de serviços, por meio de rna e modelos arima-garch. **Revista de administração mackenzie**, v. 9, n. 1, p. 130-156, 2008.

RAPOSO, R, C, T; CRUZ, A, J, O. **Stock Market Prediction Based on Fundamentalist Analysis with FuzzyNeural Networks**. 2004. Disponível em: <[http://neuro.bstu.by/ai/To-dom/My\\_research/Review-DONE/AANS-2009/Survey/3wses\\_fsfs\\_2002.pdf](http://neuro.bstu.by/ai/To-dom/My_research/Review-DONE/AANS-2009/Survey/3wses_fsfs_2002.pdf)>. Acesso em: 19 nov. 2018.

SABUHORO, J. B; LARUE, B. The Market Efficiency Hypothesis: The Case of Coffee and Cocoa Futures. **Agricultural Economics**, v.16, n.1, p. 171-184, 1997.

SILVA et.al. Finanças Comportamentais - Análise do Perfil Comportamental do Investidor e do Propenso Investidor. **Recadm**, v.7, n.2, p. 1-14, Paraná, 2008.

Silva Filho, D, M, S. **Finanças Comportamentais: o comportamento do agente decisório nos cenários de ganhos, perdas, riscos e incertezas**. Dissertação (Dissertação em economia) – UNESP. Araraquara. 2011.

SHLEIFER, A.; SUMMERS, L, H. The Noise Trader Approach to Finance. **The Journal of Economic Perspectives**, Nashville, v.4, n.2, p. 19-34, Spring 1990.

SOARES, E, R; GALDI, F, C. Relação dos modelos DuPont com o retorno das ações no mercado brasileiro. **Revista Contabilidade & Finanças**, v. 22, n.57, p. 279-298, 2011.

WEINSTEIN, N. Unrealistic Optimism about Future Life Events. **Journal of Personality and Social Psychology**, v. 39, p. 806-820, 1980.

YOSHINAGA et al. **Finanças Comportamentais: Uma Introdução**. VI I S E M E A D. USP – São Paulo, 2008.

## APÊNDICE

### APÊNDICE A: Variáveis Fundamentalistas (Input RNA)

<b>Grupo</b>	<b>Indicador</b>	<b>Código</b>	<b>Fonte</b>
Estrutura de Capital	Dívida Total Liquida	Div_Tt_Lq	Economática
Estrutura de Capital	Dívida Total Bruta	Div_Tt_Bruta	Economática
Estrutura de Capital	Dívida Bruta/Atv Tt	DivBrAtivo	Economática
Estrutura de Capital	Divida Bruta/ PL	DivBrPatr	Economática
Estrutura de Capital	Divida Liq/ PL	DLPL	Economática
Estrutura de Capital	Estr Cap (D/D+Pat Liq)	Est_Cap_PL	Economática
Estrutura de Capital	EBIT/Divida Bruta	EBITDivBr	Economática
Estrutura de Capital	EBIT/Divida Liquida	EBITDiv_Liq	Economática
Estrutura de Capital	EBIT/Desp Fin Bruta	EBITFinB	Economática
Estrutura de Capital	EBIT/Desp Fin Liquida	EBITFinL	Economática
Estrutura de Capital	Divida Bruta/Ebitda	DivBrEbitda	Economática
Estrutura de Capital	Divida Liq/ Ebitda	DivLqEbitda	Economática
Estrutura de Capital	Divida CP/ Divida Tt	DivCPDivTt	Economática
Estrutura de Capital	Exigível Tt/ Ativo Tt	ExgAt	Economática
Estrutura de Capital	Exigível Tt/ PL	ExigPL	Economática
Estrutura de Capital	Exigível Tt/Receita	ExgTtRec	Economática
Estrutura de Capital	Ativo Fixo / PL	AFxPL	Economática
Estrutura de Capital	Investimentos/PL	InvPL	Economática
Liquidez	Liquidez Geral	LiqGer	Economática
Liquidez	Liquidez Corrente	LiqCor	Economática
Liquidez	Liquidez Seca	LiqSec	Economática
Liquidez	Capital Employed	Cap_Empl	Economática
Liquidez	Capital de Giro	CapGir	Economática
Rentabilidade	EBITDA	EBITDA	Economática
Rentabilidade	LAIR + Desp Fin Liq	LAIRDFL	Economática
Rentabilidade	Giro do Ativo	GiroAt	Economática
Rentabilidade	Giro do PL	GiroPL	Economática
Rentabilidade	Margem Bruta	MrgBru	Economática
Rentabilidade	Margem Ebit (Oper)	MrgEBIT	Economática
Rentabilidade	Margem Liq	MargemLiquida	Economática
Rentabilidade	Margem Ebitda	MrgEbitda	Economática
Rentabilidade	ROA Rentab do Ativo	RentAt	Economática
Rentabilidade	ROE Rent Pat(pat final)	RenPatfim	Economática
Rentabilidade	ROE Rent Pat(pat médio)	RenPatmed	Economática
Rentabilidade	ROE Rent Pat(pat inicial)	RenPatin	Economática
Rentabilidade	ROIC (IC fim)	ROICICfim	Economática
Rentabilidade	ROIC (IC médio)	ROICIC_medio	Economática
Rentabilidade	ROIC (IC inicial)	ROICIC_inic	Economática
Rentabilidade	Invested Capital	InvestCap	Economática
Rentabilidade	Alavancagem Financeira	AlaFin	Economática
Rentabilidade	Alavancagem Operacional	AlaOpe	Economática

Dados Por Ação	Lucro por Ação	LPA	Economática
Dados Por Ação	Valor do Patr por Ação	VPA	Economática
Dados Por Ação	Vendas por Ação	VendasAcao	Economática
Dados Por Ação	EBITDA por Ação	EBITDAAcao	Economática
Indicadores de Mercado	Preco/ Lucro	PL	Economática
Indicadores de Mercado	Preço/Valor Patrim	PVPA	Economática
Indicadores de Mercado	Price Sales Ratio	PSR	Economática
Indicadores de Mercado	Preco/EBITDA	PEBITDA	Economática
Indicadores de Mercado	Preco/FLX Cx Livre	PFCF	Economática
Indicadores de Mercado	Lucro/ Preco	LP	Economática
Indicadores de Mercado	Valor de Mercado	Vlr_Mer_emp	Economática
Indicadores de Mercado	Enterprise Value Estr Cap (D/D+Valor de Merc)	EV EstCapVMerc	Economática
Indicadores de Mercado	EV/EBITDA da empresa	EVEBITDA_emp	Economática
Indicadores de Mercado	EV/EBITDA da classe	EVEBITDA_cls	Economática
Indicadores de Mercado	EV/EBIT da empresa	EVEBIT_emp	Economática
Indicadores de Mercado	EV/EBIT da classe	EVEBIT_cls	Economática
Indicadores Técnicos	TIR	TIR_media	Economática
Indicadores Técnicos	Media	Media_fechamento	Economática
Indicadores Técnicos	Div Pago por Ação	Divid_por_Acao	Economática
Indicadores Técnicos	Dividend Yield (cot fin)	DivYld_fim	Economática
Indicadores Técnicos	Dividend Yield (cot inic)	DivYld_inic	Economática
Indicadores Técnicos	Liquidez em Bolsa	Liquidez	Economática
Indicadores Técnicos	Índice de Negociabilidade	Negociabilidade	Economática
Indicadores Técnicos	Presença	Presença	Economática
Indicadores Técnicos	Volume Médio	Media_volume	Economática
Outros	Capex (Incl Invested)	Capex	Economática
Outros	Depreciação e Amortização	Depr_e_Amor	Economática
Outros	Capex (inc inv)/Deprec	CapexinvDep	Economática
Outros	Depre/Imob e Intang	DepImobInt	Economática
Outros	Ciclo Financeiro	CicFin	Economática
Outros	Ciclo Operacional	CicOpe	Economática
Outros	PM Estoque	Pmestq	Economática
Outros	PM Fornecedor	Pmform	Economática
Outros	PM Recb	Pmrecb	Economática

Fonte: Dados da pesquisa