

**INSTITUTO DOCTUM DE EDUCAÇÃO E  
TECNOLOGIA**

**FACULDADES INTEGRADAS DE CARATINGA  
CIÊNCIA DA COMPUTAÇÃO**

**EMPREGABILIDADE DA REDE NEURAL NO MERCADO DE  
COMMODITY DE CAFÉ NA BM&FBOVESPA**

**VINÍCIUS GARCIA DE OLIVEIRA**

**CARATINGA  
2016**

**VINÍCIUS GARCIA DE OLIVEIRA**

**EMPREGABILIDADE DA REDE NEURAL NO MERCADO DE  
COMMODITY DE CAFÉ NA BM&FBOVESPA**

Monografia apresentada ao Curso de  
Ciência da computação das Faculdades  
Integradas de Caratinga como requisito  
para obtenção do título de Bacharel em  
Ciência da Computação orientado pelo  
Prof. Maicon Vinícius Ribeiro.

**CARATINGA**

2016

Vinícius Garcia de Oliveira

## EMPREGABILIDADE DA REDE NEURAL NO MERCADO DE COMMODITY DE CAFÉ NA BM&FBOVESPA

Monografia submetida à Comissão examinadora designada pelo curso de Graduação em Ciência da Computação como requisito para obtenção do grau de Bacharel.



---

Prof. Maicon Vinícius Ribeiro  
Faculdades Integradas de Caratinga



---

Prof<sup>a</sup>. Msc. Glauber Costa  
Faculdades Integradas de Caratinga



---

Prof. Vagner Aquino  
Faculdades Integradas de Caratinga

*Dedico a Deus pelo dom da minha vida. Aos meus pais, João Ferreira de Oliveira e Maria das Graças Teixeira de Oliveira. A meus irmãos pelo apoio, e, em especial à minha futura esposa Bruna Aquino Ferreira, pois sem você este trabalho e muitos dos meus sonhos não se realizariam.*

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus por ter me guiado em momentos de dificuldades, pois a Ele devo tudo.

Minha família pelo apoio durante todos os momentos, em especial ao meu pai e minha mãe que me incentivaram, dando força para chegar ao fim deste grande sonho.

À minha futura esposa por seu imenso esforço, compreensão, colaboração e paciência e aos seus pais pelo carinho e apreço.

Agradeço também aos meus professores, que contribuíram para minha formação acadêmica, compartilhando seus conhecimentos e experiência de vida, que foram de grande valia para minha formação.

Agradeço também a todos envolvidos neste trabalho que colaboraram para seu desenvolvimento direta ou indiretamente.

*Porque onde estiver o vosso tesouro, aí estará também o vosso coração.*

*Mateus 6:21*

## RESUMO

Prever o Mercado de Ações tem sido objeto de estudos por diversas áreas, por possuir um alto índice de volatilidade e complexidade. Fato este que leva o investidor a especular, analisar, buscar previsões, estudar ações ou commodities, com propósitos diversos em apostar altas ou baixas, mas com objetivo em obter ganho.

O Brasil (BR) é um dos maiores exportadores de café arábica do mundo, e, entre seus estados, Minas Gerais (MG) é o responsável por grande parte da produção, que é negociada na bolsa de Mercado Futuro como BM&FBovespa.

Observando esta demanda por ferramentas que buscam realizar a previsão dos mercados futuros, o presente trabalho teve por objetivo, analisar dados históricos das previsões diárias de café futuro, empregando o uso de redes neurais com o *Multi Layer Perceptron* com algoritmo *backpropagation*, para classificação das previsões diárias de café futuro.

Os resultados e avaliações da Rede Neural são apresentadas por meio de relatórios gráficos, que demonstraram o desempenho alcançado com os treinamentos e classificações, junto às previsões diárias das cotações de café futuro que apresentaram resultados satisfatórios.

**Palavras-Chaves:** Mercado Futuro, *Commodity*, Café, Inteligência Artificial, *Multilayer Perceptron*.

## LISTA DE ILUSTRAÇÕES

FIGURA 1 - Estrutura de um neurônio biológico. (CANTO, 2009). .....	26
FIGURA 2 - Modelo de um neurônio artificial.....	27
FIGURA 3 - Função não linear (SIMEÓN, 2015).....	27
FIGURA 4 - Logo tipo do WEKA .....	29
FIGURA 5 - Pesquisa por cotações pregão online BM&FBovespa .....	32
FIGURA 6 - Cotações café para cotações de operação com contratos ICFZ2016 ...	33
FIGURA 7- Acesso área de serviços e históricos.....	33
FIGURA 8 - Acesso área derivativos pesquisa por pregão. ....	34
FIGURA 9 - Pesquisa por pregão.....	34
FIGURA 10 - Estrutura arquivo diário.....	35
FIGURA 11 - Planilha descrição dos campos. ....	36
FIGURA 12 - Treinamento e classificação pela RNA para cotações diárias. ....	42
FIGURA 13 - Treinamento e classificação pela RNA para ajustes diários. ....	43
FIGURA 14 - Gráfico com a previsão para cinco próximos dias vezes a cotação real. .....	45
GRÁFICO 1 - Principais países produtores, volume em mil sacas de 60 Kg, anual. 16	
GRÁFICO 2 - Principais países consumidores, volume em mil sacas de 60 Kg anuais.....	17

## LISTA DE TABELAS

TABELA 1 - Parâmetros usados para configuração da MLP para realizar treinamento da rede e definir .....	38
TABELA 2 - Configurações para RNA.....	38
TABELA 3 - Melhores RNA para camadas ocultas. ....	39
TABELA 4 – Melhores RNA com valores para taxa de aprendizado.....	40
TABELA 5 - Melhores RNA com valores para ajustes de pesos. ....	41
TABELA 6 - Melhores RNA nos treinamentos.....	41
TABELA 7 - Valores mínimos e máximo de 5 dias antes da previsão.....	44
TABELA 8 - Previsão para os 24 a 28 de novembro de 2016. ....	45

## **LISTA DE SIGLAS**

API - Application Programming Interface

BM&F - Bolsa de Mercadorias e Futuros

BM&FBovespa - Bolsa de Valores, Mercadorias e Futuros de São Paulo

Cecafe - Conselho dos Exportadores de Café do Brasil

DOC - Departmente of Conservatio

EUA - Estados Unidos da América

GPL - General Public License

Ibovespa - Índice Bovespa

ICB - Commodities Brasil

ICO - INTERNATIONAL COFFEE ORGANIZATION

KDD - Knowledge Discovery in Databases

LRQ - Learning Vector Quantisation

MG - Minas Gerais

MLP - Multilayer Perceptron

PHP - Hypertext Preprocessor

RNA - Rede Neural Artificial

SGBD - Sistema de Gerenciamento de Bando de Dados

SP - São Paulo

USDA - United States Department of Agriculture

WEKA - Waikato Environment for Knowledge Analysis

## SUMÁRIO

INTRODUÇÃO .....	13
1 REFERENCIAL TEÓRICO .....	15
1.1 MERCADO E PRODUÇÃO DO CAFÉ .....	15
1.1.1 O Mercado de café Mundo .....	15
1.1.2 O café no Brasil .....	17
1.1.3 Produção de café em Minas Gerais .....	18
1.2 MERCADOS DE AÇÕES NA BM&FBOVESPA .....	19
1.2.1 BM&FBovespa .....	20
1.2.2 Ações .....	20
1.2.3 Mercado Futuro .....	21
1.2.4 <i>Commodities</i> .....	22
1.2.5 Contratos Futuros de café .....	23
1.3 DADOS HISTÓRICOS DE COTAÇÕES .....	24
1.4. REDES NEURAIS ARTIFICIAIS .....	24
1.4.1 O neurônio Biológico .....	25
1.4.2 Neurônio artificial.....	26
1.4.3 Rede <i>Multilayer Perceptron</i> .....	28
1.4.4 O Algoritmo <i>Backpropagation</i> .....	28
1.4.5 WEKA.....	29
2 METODOLOGIA.....	31
2.1 OBTENÇÃO DOS DADOS.....	32
2.2 MINERAÇÕES DOS DADOS.....	35
2.3 SELEÇÃO DOS ATRIBUTOS .....	37
2.4 CONFIGURAÇÕES PARA REDE MLP.....	37
2.5 TREINAMENTOS E ANÁLISES DOS TESTES .....	39

2.6 CLASSIFICAÇÃO E VALIDAÇÃO DA REDE .....	42
2.7 CLASSIFICAÇÃO DAS PREVISÕES COM RNA.....	44
3 CONCLUSÕES .....	46
TRABALHOS FUTUROS .....	47
REFERÊNCIAS.....	48

## INTRODUÇÃO

A produção e comercialização de café teve grande importância para a economia brasileira entre os séculos XIX e XX, representando em certo momento, cerca de 70% da produção mundial de café (BAPTISTA, 2015). Com a crise de 1929, que atingiu a economia americana e a do mundo, houveram fortes reflexos no setor cafeeiro. Fato que, desde então fez o café perder espaço para outros cultivos (SERRA, 2010). Mesmo perante a perda do café para outras áreas, este ainda se encontra como a terceira *commodity* mais negociada, com cerca de 12.655 contratos sendo realizados, pelo balanço anual de 2015 feito pela BM&FBovespa (BM&FBOVESPA, 2015).

O risco do investimento no mercado de ações é eminente, sendo alvo de estudos por diversos economistas, órgãos e instituições financeiras (SILVA, M., 2014). Diante desta situação, os Mercados Futuros são a forma mais efetiva para medir as oscilações de preços, gerando a possibilidade de minimizar os riscos nos investimentos (MACIEL, 2016).

Estes riscos associados a outros fatores que possibilitam a crise são resultantes da soma de fatores econômicos e sociais que são vivenciados em todo o mundo. Fatores que por sua vez influenciam diretamente os investidores, para realizarem a tomada de decisões ao efetuarem a compra ou venda de um determinado ativo, sendo levada em conta sua experiência e conhecimento no mercado (NICHI, 2016).

As técnicas computacionais e modelos matemáticos como as redes neurais têm demonstrado bons resultados nas previsões em aplicações que envolvem séries temporais, como no mercado de ações que possui uma alta complexidade na formação das cotações (TEIXEIRA, 2012).

O avanço produzido no campo das Redes Neurais Artificiais (RNA) tem elevado as chances de ganho no mercado financeiro, proporcionando investimentos mais sofisticados, que geram um aumento consequente no nível de risco (MARTINS, METTE e MACEDO, 2008).

Assim como os autores anteriormente citados que obtiveram bons resultados em suas análises e aplicações com o uso das RNA, o presente trabalho busca nas sessões subsequentes, realizar um levantamento referencial sobre o mercado de *commodity* de café na BM&FBovespa por meio da elaboração conceitual, formação das RNA e uso do algoritmo *Multilayer Perceptron* (MLP), descrevendo também todo o processo realizado para o desenvolvimento deste estudo na previsão de mercados futuros diários de café, analisando as cotações diárias da BM&FBovespa no período de 2003 a 2016, descrevendo os testes, análises e resultados obtidos.

## 1 REFERENCIAL TEÓRICO

Neste referencial teórico são apresentados os conceitos fundamentais para entendimento do processo e abordagem, que são empregados ao decorrer do presente estudo. Subdividido em tópicos, esse, apresenta os principais mercados consumidores, produção e cenário Brasileiro de café.

### 1.1 MERCADO E PRODUÇÃO DO CAFÉ

“O café é uma das bebidas mais consumidas em todo mundo” (BOLINI, 2008), fato este que faz o café ser tão apreciado. Não somente por sua propriedade nutricional ou valores calóricos, mais pelo seu sabor, sensação e aroma. Além do prazer em estimar uma bebida tão nobre que pode ser encontrada e preparada mundo afora (ZAIDAN, 2015).

Segundo informações divulgadas pelo Conselho dos Exportadores de Café do Brasil (Cecafe), em setembro de 2016 o País se classificou como o maior exportador no mercado mundial, com exportações entres safras 2015/2016 em torno de três milhões de sacas aproximadamente.

#### 1.1.1 O Mercado de café Mundo

O mercado e consumo de café no mundo vem sofrendo queda ao longo das safras, estando em destaque uma queda de 17% em relação a última safra mundial que foi estimada em 151,3 milhões segundo o *INTERNATIONAL COFFEE ORGANIZATION* (ICO). Este é um importante órgão responsável pelo Acordo Internacional do Café (AIC), composto por diversos membros e governos que correspondem a 94% de toda a produção e cerca de 75% do consumo mundial além de fornecer importantes dados estatísticos mundiais sobre as safras e o consumo.

O Brasil está entre os maiores produtores de café em nível mundial se destacando em primeiro lugar por número de sacas vendidas anualmente (43,2 milhões). Em seguida estão o Vietnã (27,5 milhões), Colômbia (14,0 milhões), Indonésia (12,3 milhões), Etiópia (6,7 milhões), Índia (5,8 milhões), Honduras (5,4 milhões), Uganda (4 milhões), Guatemala (3,4 milhões) e Peru (3,3 milhões), como observado no Gráfico 1 a seguir.

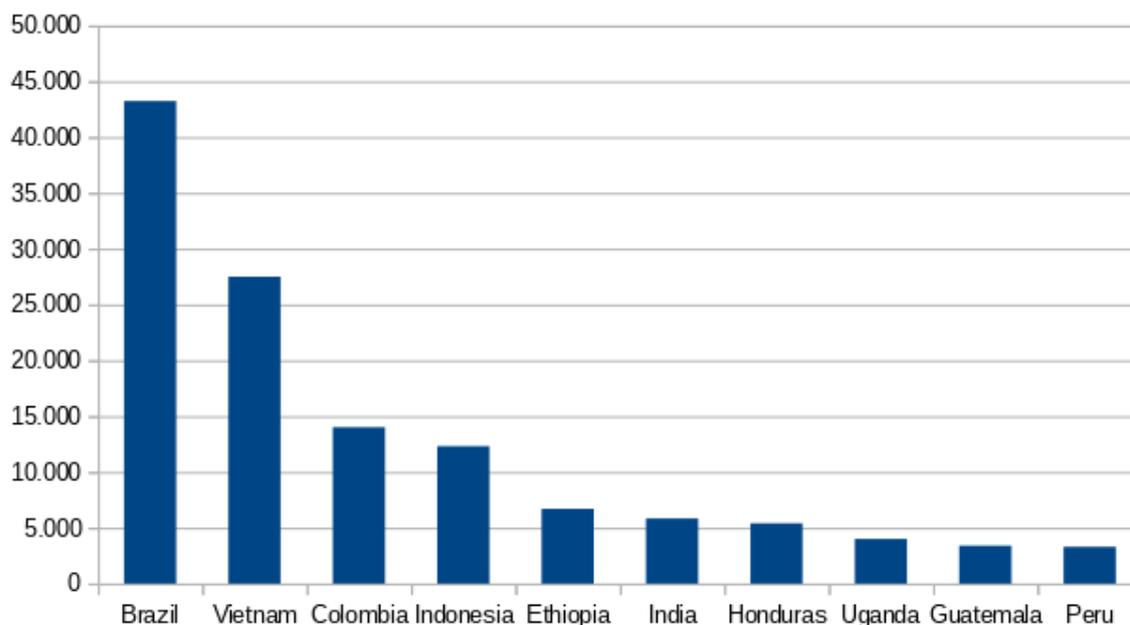


GRÁFICO 1 - Principais países produtores, volume em mil sacas de 60 Kg, anual.

Fonte: OIC, 2016.

Além disso, o Brasil se destaca também por ser o segundo maior consumidor do produto (20,5 milhões ao ano), ficando atrás dos Estados Unidos da América (EUA) (24,4 milhões), seguidos por Japão (7,6 milhões), Indonésia (4,5 milhões), Federação Russa (3,8 milhões), Etiópia (3,7 milhões), Filipinas (3,0 milhões), México (2,35 milhões), Vietnã (2,3 milhões) e Índia (2,2 milhões) como ilustrada a seguir.

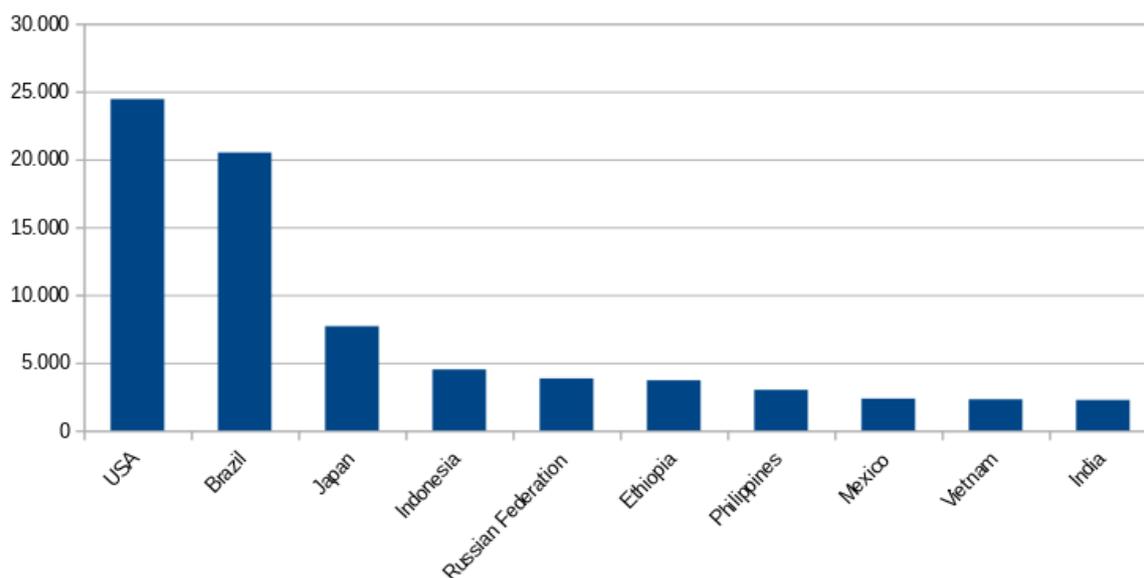


GRÁFICO 2 - Principais países consumidores, volume em mil sacas de 60 Kg anuais.

Fonte: OIC, 2016.

O Gráfico 2 destaca uma grande margem de consumo no mercado Americano e no próprio mercado interno Brasileiro em relação ao mercado Japonês e aos demais.

### 1.1.2 O café no Brasil

Para compreendermos a atual posição do café em nosso País, devemos recordar sua origem. Segundo Rocha (2009) a “Etiópia é, possivelmente, o exportador de café mais antigo do mundo [...]”, tendo sua origem de Kaffa na Absínia (antiga Etiópia), que ao longo do tempo, foi se espalhando para Europa no século XIV. A palavra café, assim como grão, tem sua origem árabe na palavra *qahwa*, que quer dizer vinho, e chegou a ser conhecido na Europa como o vinho da Arábia (PAGIATTO, 2013).

A partir da Europa, o cultivo do café foi difundido no mundo pelos colonizadores europeus e chegou ao Brasil em 1727, pela Guiana Holandesa. Seu cultivo se iniciou no Norte do país (primeiro em

Belém e em seguida no Maranhão) e, a partir dessa região, se difundiu pelo país até atingir a região do Vale do Paraíba e interior paulista, onde se tornou a principal atividade no século XIX (BAPTISTA, 2015).

Após a chegada do café aos demais estados como Minas Gerais (MG) e São Paulo (SP), não levou muito tempo para se tornar o pilar da economia brasileira, por quase um século, chegando então à crise de 1929, na qual surgiu o pior cenário para o setor cafeeiro brasileiro, que ao término do colapso voltou a ter grande importância econômica nas exportações (GONÇALVES, 2014).

No cenário atual mundial, vivencia quedas na produção e no consumo. Porém, mesmo diante de tal problemática, o BR mantém seu posto de maior produtor como nos informam os dados divulgados pela *United States Department of Agriculture* (USDA) em 2015. A mesma afirma que o país ocupa a primeira posição em exportação da *commodity* por vários anos consecutivos. Mesmo com a tendência de baixo consumo da bebida o BR segundo a ABIC atingiu um recorde na exportação, entre as safras de 2014/15 (julho a junho) com 36,49 milhões de sacas. Um aumento considerado de 6,9% em relação ao ano anterior (ABIC, 2016).

Produzindo os mais variados tipos de café em que se destacam as espécies *Coffea arabica* (Café arábica) e *Coffea conphora* (Café conilon ou Café robusta). O café arábica se realça entre os demais, por conservar seu aroma e sabor peculiar mesmo após o processo de torrefação, elevando sua classificação. Já o conilon é geralmente usado para formação “*blends*” (ligas de cafés) por ser de qualidade inferior, e para montar ligas com o arábica inferior objetivando melhoria no corpo da bebida elevando o seu preço de venda (DUARTE, 2015).

### 1.1.3 Produção de café em Minas Gerais

O Café no cenário mineiro divide-se em quatro grupos. Em primeiro vem as regiões Sul e Centro-Oeste, responsáveis por cerca de 52,58% (523.506,0) da produção por hectares seguidas pela Zona da Mata, Rio Doce e Central com 26,22%

(261.109,0), logo após, o Triângulo, Alto Paranaíba e Noroeste com 18,39% e em último lugar o Norte, Jequitinhonha e Mucuri com 2,81% (28.024,0). Juntas, estas áreas correspondem a 73,22% (995.715,0) da produção em hectares nacional, de acordo com estimativas para o mês de setembro de 2016 (Conab, 2016).

Por apresentar relevo acidentado, com altitude média superior a 650 metros, a ocorrência dos ventos catabáticos e anabáticos são comuns na região contribuindo para a ocorrência de microclimas entre as montanhas extremamente favoráveis ao cultivo de café. A ação constante dos ventos, associado a presença da vegetação entre os vales e montanhas, contribuem para o clima ameno na região, também denominado como clima de montanha (ZAIDAN, 2015).

Dessa forma, o clima montanhoso encontrado nas regiões de MG torna favorável a produção de café de boa qualidade, como a linha de cafés especiais que com sabor único e restrito ao seu potencial alcançam classificação valiosa (LUZ, 2014).

## **1.2 MERCADOS DE AÇÕES NA BM&FBOVESPA**

A composição do Mercado de Ações, também conhecido por Mercado de Capital, tomado por base no Sistema Financeiro, compreende recursos de oferta e demanda além de ser constituído por dois tipos de agentes. Os Econômicos e Superavitários. Os primeiros são pessoas, empresas, ou o Estado. O segundo são aqueles que em geral consomem menos recurso do que ganham desta forma procuram poupá-lo. Por outro lado, alguns agentes precisam de mais recursos do que possuem tendo de recorrer à poupança, ficando conhecidos como agentes deficitários (Mercado de Valores Mobiliários Brasileiro, 2014).

### 1.2.1 BM&FBovespa

O surgimento da BOVESPA ocorreu por volta de 1890 a partir da iniciativa de um grupo de empresários paulistas em criar Bolsa Livre de Valores. Posteriormente a mesma montou a Bolsa Mercantil e de Futuros, pouco tempo depois ela passou a incorporar a Bolsa de Mercadorias de São Paulo, denominando-se como Bolsa de Mercadorias e Futuros (BM&F). A BOVESPA e BM&F juntas realizaram importantes avanços tecnológicos, onde um deles abriu portas para pequenos investidores realizarem operações através dos “*home broker*”, sistema online de pregão onde o investidor pode realizar ordem de compra ou venda de ações. Chegando em 2008 onde com o objetivo de aumentar expressivamente o mercado acionário brasileiro, as duas bolsas se uniram, para formar Bolsa de Valores, Mercadorias e Futuros de São Paulo (BM&FBovespa), tornando uma das maiores bolsas de valores e ações no mercado e a segunda maior das Américas (KEMPER, 2012).

### 1.2.2 Ações

Uma ação pode ser denominada como título, ela é a representação de uma fração do capital social da empresa emissora, onde seus títulos são negociados por intermédio da bolsa de valores, possibilitando que o investidor obtenha ganhos do patrimônio de uma empresa (BASILEIRO, 2013).

Estes títulos são sempre negociados em uma bolsa, no caso em questão, na BM&FBovespa, onde cada investidor sempre busca por meio da aquisição dos mesmos, obter ganhos superiores à sua compra, expressando assim o sucesso de uma empresa em que ele adquiriu partes (BORGES, RULLO, et al. 2011).

O Mercado de Ações proporciona giro de capital entre poupança e agentes econômicos, criando um círculo virtuoso entre estes recursos que produzem investimentos e crescimento econômico, contribuindo significativamente para o aumento de renda e avanço dos países mais desenvolvidos que possuem mercados mais dinâmicos, diversificado e fortalecido (QUIRINO, 2012).

### 1.2.3 Mercado Futuro

De acordo com Almeida (2011) e Neto (2008) existem quatro tipos de mercados que operam diversas negociações e *commodities* agrícolas. São eles: Mercado Físico ou Mercado *Spot*, Mercado *Termo*, Mercado de *Opções* e Mercado *Futuro*.

O Mercado Físico ou Mercado *Spot* é conhecido pelas características de suas transações imediatas que consente nos pagamentos feitos à vista, e mercadorias sendo entregues ou transferidas em prazo muito curto, quando não no mesmo dia, estes tipos de operações podem ser muito utilizadas por agricultores que necessitam de capital emergente, ou na venda de gás e energia elétrica, que podem ser entregues em curtos prazos, que assim estabelecem as características que compõem o tipo de mercado.

O Mercado *Termo* se caracteriza pela negociação direta entre vendedores e compradores, estabelecendo o preço e data de entrega da *commodity* ou ação acordada entre os negociantes para uma data futura. Ao fim do prazo determinado entre as partes, o produtor realiza a entrega da *commodity* dentro dos termos acordados e o comprador realiza o pagamento. A obtenção de ganho entre as partes está no momento de liquidez do acordo, na qual o preço acordado pode estar abaixo do mercado, gerando assim ganho para o produtor que já esperava por esta baixa, ou acima do preço combinado indicando lucro para o comprador.

O Mercado de *Opções* negocia o direito de compra ou venda em contratos com um preço fixo para uma data futura. Para que o comprador conhecido como titular tenha o direito ao contrato é necessário o pagamento de um valor prêmio, valor que é pago ao vendedor também conhecido com lançador, dando assim direitos ao titular de exercer o curso das negociações entre os contratos no mercado, ficando o lançador na obrigação de executar as ações definidas pelo titular, para os rumos do contrato.

O Mercado *Futuro* é uma evolução do mercado termo, com várias semelhanças, porém o diferencial entre eles está na forma de liquidação que ocorre diariamente ao término de cada pregão onde são calculados os lucros ou perdas, para as partes do contrato, que agregam margens mínimas de chances para garantir

o preço de produção, estabilizando o risco no mercado físico, que oscilam resultante as especulações dos mercados.

A diferença entre os preços à vista e no mercado futuro é denominada de “base” ou “prêmio”, sendo que uma base pequena indica que os dois preços caminham juntos e que as expectativas do mercado apontam para a manutenção da tendência dos preços no mercado à vista. Quando o mercado acredita que o movimento de alta vai continuar, os preços futuros são iguais ou mais altos que os preços à vista. Porém, períodos de variação da base são muito frequentes e estão associados à dinâmica dos mercados futuros, altamente especulativa. Sobre essa dinâmica e os determinantes das variações da base, que dependem da natureza do ativo subjacente — mercadorias estocáveis, não estocáveis, ativos financeiros e divisas. Farhi (1998).

Para Torres (2005) “[...] mercados futuros possuem sua origem baseada na necessidade de transferência dos riscos atrelados à produção e comercialização de produtos agrícolas.”.

#### **1.2.4 Commodities**

Uma *commodity* (plural *commodities*) é uma palavra inglesa utilizada para descrever uma mercadoria que se encontra em seu estado bruto ou com pequeno nível de industrialização, comercializada em grandes quantidades de forma uniforme (JALE, 2015).

*Commodities* é uma maneira de descrever produtos em sua forma bruta, ou com baixo índice de industrialização. Comercializado em grande escala, tendo diversos fornecedores, com qualidade homogênea, preço variável independente do fornecedor ou produtor, estabelecendo assim um mercado competitivo (PUPGA, 2008).

O pesquisador Romeu (2014) descreve *commodities* como:

Mercadorias que apresentam uma característica fundamental: são bens fungíveis, equivalentes em termos de especificações estabelecidas independentemente de quem as produz. O seu preço à vista é determinado pelo mercado (*market price*) e suas flutuações ocorrem em função da oferta e da demanda, conforme definições da economia clássica (ROMEY, 2014).

O Mercado de *commodities* é subdividido em quatro tipos:

- *Commodities* minerais: carvão, gás natural, minério de ferro e petróleo;
- *Commodities* financeiras: real, euro, dólar, entre outras moedas;
- *Commodities* ambientais: água, madeira, energia, etc;
- *Commodities* agrícolas: arroz, milho, trigo, cevada, café entre outros.

Para desenvolvimento desta pesquisa será utilizada a *commodities* agrícola, em específico as *commodities* de café em contratos futuros. A fim de prever mercados pósteros da economia cafeicultora.

### 1.2.5 Contratos Futuros de café

Segundo a Associação Brasileira da Indústria Cafeicultora (ABIC), contratos futuros de café Arábica cru tipo 4-25 (4/25) sendo a melhor bebida dura, cuja classificação é feita em amostra de 300 gramas de café beneficiado, que contenha cerca de 36 defeitos e pontuação de 25.

Cada contrato é composto por 100 sacas, que pesam 60 kg a unidade. Estes por sua vez são cotados e negociados no mercado futuro e sua liquidação é realizada em Dólares.

Todas as negociações são realizadas via *Home Broker* que é o Sistema Eletrônico de Negociações, fornecido por uma corretora, que mantém vínculo com estas permitindo que cliente realize compra e venda de Ativos. Segundo o Manual de Procedimentos Operações Ações BM&FBovespa 2013.

Segundo BM&FBovespas a formação dos códigos de identificação dos contratos futuros de café começa pelo código ICF precedidos pela letra do mês

vencimento do contrato. A identificação de cada letra por mês é dada da seguinte forma: Janeiro (F), Fevereiro (G), Março (H), Abril (J), Maio (K), Junho (M), Julho (N), Agosto (Q), Setembro (U), Outubro (V), Novembro (X) e Dezembro (Z). Os principais meses de vencimento dos contratos de café futuro são Março, Maio, Julho, Setembro e Dezembro. E o último sufixo que forma o código é o ano do contrato, sendo que um contrato com vencimento para dezembro de 2016 tem o seguinte código ICFZ16. Esta lógica é empregada para identificar dentre as muitas cotações históricas e futuras.

### **1.3 DADOS HISTÓRICOS DE COTAÇÕES**

Os dados históricos ou Séries históricas de cotações são disponibilizados diariamente pela BM&FBovespa por meio de seu portal na internet, contendo informações desde o ano de 1986, que são informações que levam em consideração a forma da moeda em cada período, não sendo realizado nenhum ajuste inflacionário sobre as informações (BM&FBovespa, 2016).

Todas as informações disponíveis no *website* podem ser usadas para fins pessoais e estudos. O uso das informações para comercialização, locação ou distribuição são restritos. Para acessá-las torna-se necessário a posse do termo de autorização cedido pela BM&FBovespa.

Petry (2010) e Rambo (2014) por meio de dados históricos notaram que o desempenho de ações tendem a repetição de cotações. Compreender e identificar a forma como estas acontecem fazem parte da análise técnica com o uso de cálculos matemáticos, estatísticos e de RNA.

### **1.4. REDES NEURAIS ARTIFICIAIS**

Os estudos sobre RNA tiveram início significativo após apresentação do primeiro modelo algébrico que descreveu o funcionamento do neurônio artificial,

proposto por McCulloch e Pitts (1943). Este recebeu o nome de MCP em condecoração aos autores (SILVA, G., N. 2014).

Após os primeiros estudos propostos por McCulloch o psicólogo canadense Donald Hebb escreveu em seu livro "*The Organization of Behavior*" (1943), a primeira ideia de progressão de pesos sinápticos para os neurônios, não muito tempo depois Roseblatt (1958) desenvolveu o primeiro neurônio artificial *perceptron* e o treinamento supervisionado (SIMEÓN, 2015).

RNAs é a definição usada para representar cálculos matemáticos e algoritmos computacionais. Busca reproduzir características do processamento e reconhecimento de padrões como o do cérebro humano. Esta pode ser descrita como uma caixa de processamento (NETO, 2016).

Uma rede neural é composta por vários neurônios interligados entre si, além de manterem peso de conexão. Cada um possui entradas, que juntas são analisadas e por meio de um processamento ou pela função de ativação geram uma saída (VALE, 2016).

#### **1.4.1 O neurônio Biológico**

Para compreender o funcionamento de um neurônio artificial, primeiramente devemos entender o neurônio biológico, para que desta forma entendamos a similaridade existente entre eles.

De uma forma básica podemos fragmentar o neurônio humano em três partes. Este é constituído pelo corpo celular que recebe a informação transmitida por meio de seus dendritos estimulados pelo ambiente externo. Estes estímulos por sua vez são processados e enviados para o axônio que a seguir irá transmitir a mensagem. Esta divisão pode ser observada na Figura 1 (NETO, 2014).

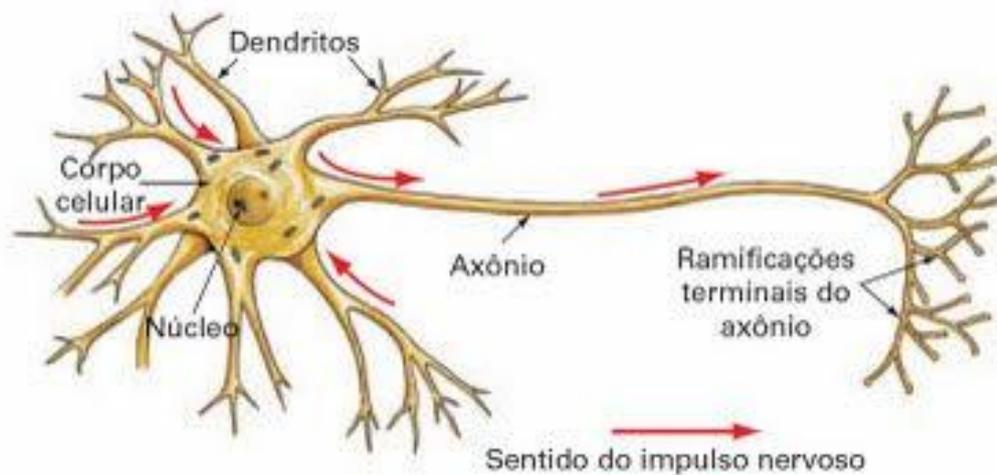


FIGURA 1 - Estrutura de um neurônio biológico. (CANTO, 2009).

Segundo Ferro (2013), as redes neurais são formadas por neurônios interconectados, ou seja, as sinapses, dão forma ao pensamento, sentimento, empatia e a cognição. É por intermédio da sinapse que um neurônio se conecta com até 10.000 outros. As interconexões podem chegar a um total de 100 bilhões dentro do nosso cérebro.

#### 1.4.2 Neurônio artificial

Bocato (2013), descreve o neurônio como a unidade-base para o processamento de incentivos provenientes de outro neurônio ou do meio externo. Estes neurônios geram as conexões entre saída e entrada descritas como sinapses, assim como no neurônio biológico anteriormente citado. A seguir na Figura 2 pode ser visto um modelo de neurônio artificial.

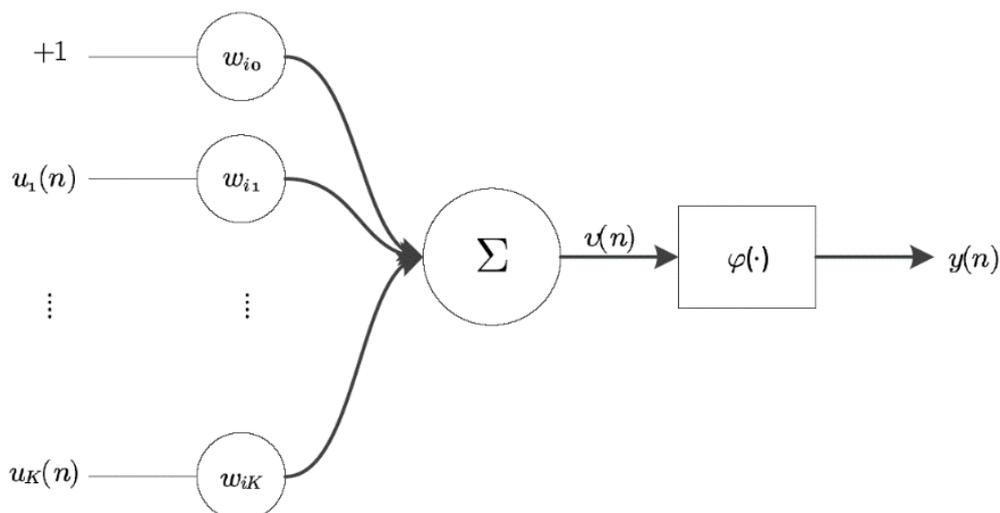


FIGURA 2 - Modelo de um neurônio artificial  
 Fonte: BOCCATO, 2013, p. 8.

Um neurônio artificial é responsável por processar as inúmeras informações e decidir a relevância da informação gerando uma saída que são tomadas como base para ativação linear, atribuindo pesos para cada sinapse onde tende a aumentar ou invalidar a entrada para o posterior (FRANCO, 2015).

$$y(n) = \varphi \left( \sum_{k=1}^K w_{ik} u_k(n) + w_{i0} \right)$$

FIGURA 3 - Função não linear (SIMEÓN, 2015).

Para descrever o cálculo que é realizado pela função linear, a Figura 3 nos mostra o processo matemático para chegar a saída  $y(n)$  que fornece a informação a ser repassada para o próximo neurônio artificial. Fazendo a soma ponderada entre os sinais da entrada  $u_k(n)$ ,  $k = 1, 2, \dots, K$ , junto com a multiplicação dos pesos sinápticos de  $w_{ik}$ , o resultado da operação é avaliado por uma função não linear  $\varphi$  (SIMEÓN, 2015).

### 1.4.3 Rede *Multilayer Perceptron*

A rede *Multilayer Perceptron* (MLP) ou rede multicamadas são redes com capacidades para resolução de problemas complexos, pois, sua organização compreende em neurônios que estão dispostos paralelamente entre si (GIACOMEL, 2016).

A MLP é conhecida também pelo tipo de método de trabalho. Nesta, a propagação da mensagem é repassada pelo *feedforward* que se dá somente em um sentido. Na medida em que as informações são recebidas por meio das camadas de entrada os sinais serão processados e levados para as camadas de saída (SOUZA, 2012).

A rede MLP combinada ao algoritmo *Backpropagation* tem difundindo de forma impactante no mundo das RNAs pelo seu método de retro-propagação (LEMES, 2014).

### 1.4.4 O Algoritmo *Backpropagation*

O *Backpropagation* é o algoritmo mais aplicado nas redes MLP, perante tal informação, pode-se afirmar que na “[..] maioria dos algoritmos de aprendizagem das redes multicamadas utilizam como base este algoritmo *Backpropagation* e este foi principal responsável pelo retorno do interesse de pesquisadores nos estudos das RNA” (VALE, 2016).

O algoritmo de retropropagação opera em dois passos. No primeiro passo, é apresentado um conjunto de dados na camada de entrada da rede, percorrendo-a camada por camada, até que a resposta seja produzida pela camada de saída. Neste passo, nenhum peso sináptico é alterado. No segundo passo, a saída obtida é comparada a saída desejada. Caso a saída não estiver correta, um erro é calculado a partir da saída até a camada de entrada, modificando os pesos das conexões conforme o erro é retropropagado (KRIEGER, 2012, p.44).

O algoritmo MLP com *Backpropagation* pode ser encontrado na biblioteca do *Waikato Environment for Knowledge Analysis* (WEKA) assim como outros algoritmos utilizados nas RNAs.

#### 1.4.5 WEKA

O WEKA nasceu a partir da iniciativa de pesquisadores na Universidade de Waikato na Nova Zelândia como uma ferramenta para mineração de dados, que busca prover conhecimento por meio de análises realizadas, sendo o termo usado *Knowledge Discovery in Databases* (KDD) (STULP, 2014).

Segundo Araújo (2009) o WEKA implementa uma série de algoritmos e ferramentas para análise e processamento de dados, sendo distribuído e licenciado como software livre pela *General Public License* (GPL), para que muitos possam estudar e contribuir para seu desenvolvimento. A linguagem de programação utilizada pelo WEKA e pelos algoritmos, bibliotecas e interface são implementadas com o JAVA.

O uso do nome WEKA e sua logo Figura 4, se dá por causa de uma ave que vive nos arredores das ilhas da Nova Zelândia, onde se situa a Universidade de Waikato (NOGUEIRA, 2013).



FIGURA 4 - Logo tipo do WEKA

Fonte: Site waikato

A ave *Apteryx* popularmente conhecida como kiwi que é o símbolo do WEKA se encontra em extinção, porém com iniciativas *Departmente of Conservatio* (DOC)

da Nova Zelândia junto com outros institutos tem criado reservas para aumentar a população de kiwi (DOC, 2016).

## 2 METODOLOGIA

Quando um investidor pensa em realizar a compra de uma *commodity* como o café, ele logo observa diversas variáveis e tendências de mercado com intuito de em um futuro próximo, realizar a venda e obter um lucro elevado sobre o montante investido. Porém, para investidores com pouca ou nenhuma experiência no mercado financeiro, dificilmente o investidor alcançará bons ganhos por conta da alta complexidade e volatilidade que envolve o mercado.

Levando em conta a atual localidade da instituição e suas proximidades onde as principais atividades agrícolas são cultivos e manejos de cafezais que proporciona geração de emprego e renda para população local aquecendo e movimentando a economia regional.

Assim como os investidores almeja ganhos em seus investimentos com o produtor agrícola também não diferente. Com o mesmo conceito que os investidores o produto visa obter o melhor ganho possível com a venda de seu produto.

O presente trabalho toma como propósito analisar cotações históricas sobre *commodity* de café, mais em específico os contratos futuros que são negociados diariamente na BM&FBovespa, aplicando técnicas de mineração de dados, subdividindo as informações encontradas em dois universos, o primeiro para treinamento da rede MLP que implementa o algoritmo *backpropagation* na aprendizagem e o segundo conjunto de informações para a classificação e validação da rede neural, que posteriormente é utilizada na classificação e autenticar as previsões geradas aleatoriamente dentro de um período estabelecido com as cotações reais. Para início do estudo é necessário obter as cotações de café futuro, com esta finalidade a sessão seguinte descreve em detalhes o processo abordado para obter os dados.

## 2.1 OBTENÇÃO DOS DADOS

Todas as transações que acontecem do mercado financeiro brasileiro na BM&FBovespa, estão disponíveis para quaisquer empresas e pessoas que queiram fazer usos para fins próprios, tais informações podem ser visualizadas por meio de ferramentas *home broker* que são disponibilizadas para pessoas que sejam associadas a uma corretora ou através do site da BM&FBovespa que conta um painel com pregão eletrônico como pode ser observado na Figura 5.

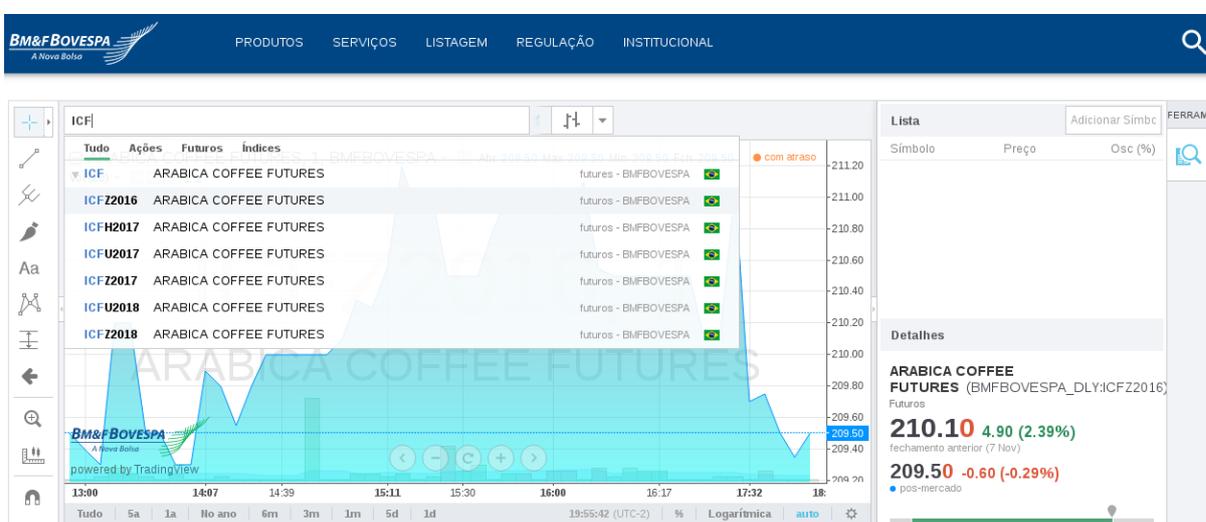


FIGURA 5 - Pesquisa por cotações pregão online BM&FBovespa

Fonte: BM&FBovespa 2016.

Como mostra a Figura 5 anteriormente é necessário ter um prévio conhecimento sobre os tipos de códigos existentes, como foi abordado no tópico 1.2.5 Contratos Futuros de café. Com o uso da ferramenta disponível pela BM&FBovespa pôde ser vista todas as cotações de um contrato ou ação selecionada como mostra a Figura 6 posteriormente.

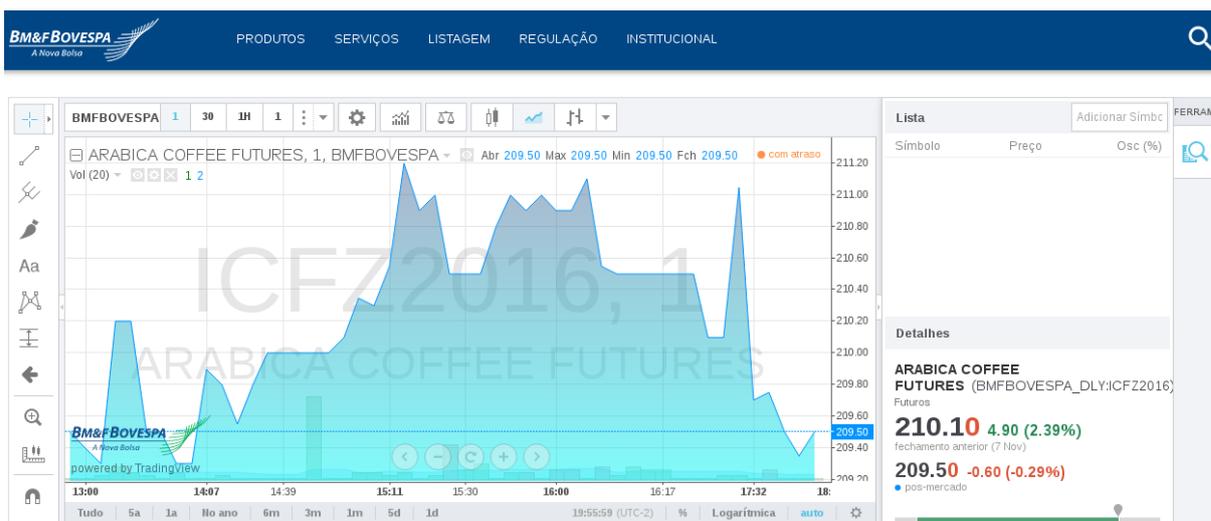


FIGURA 6 - Cotações café para cotações de operação com contratos ICFZ2016

Fonte: BM&FBovespa 2016.

Na Figura 6 foi selecionado um contrato futuro de café com vencimento para dezembro de 2016 com o seguinte código ICFZ2016. Com base na ferramenta que é acessível no *website* da BM&FBovespa futuro foi utilizado como parâmetro para certificação da integridade da informação após passar pelo processo que envolve a importação dos dados para um banco de dados MySQL.

Para obter as cotações relacionadas as *commodities* é necessário acessar o site da BM&FBovespa, passar o *mouse* sobre o menu de “Serviços” e clicar no submenu de “Histórico” como exhibe a Figura 7.

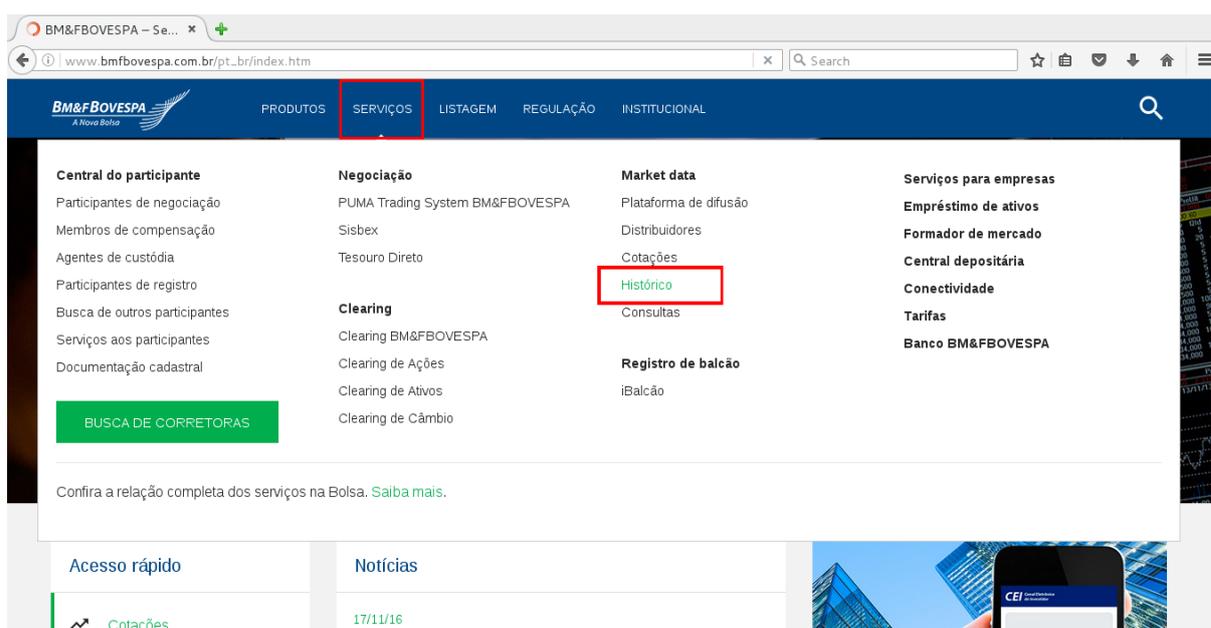


FIGURA 7- Acesso área de serviços e históricos.

Fonte: BM&FBovespa 2016.

Será carregada uma página semelhante à Figura 8, na área de Derivativos, sendo necessário clicar em pesquisa por pregão que se encontra em destaque.

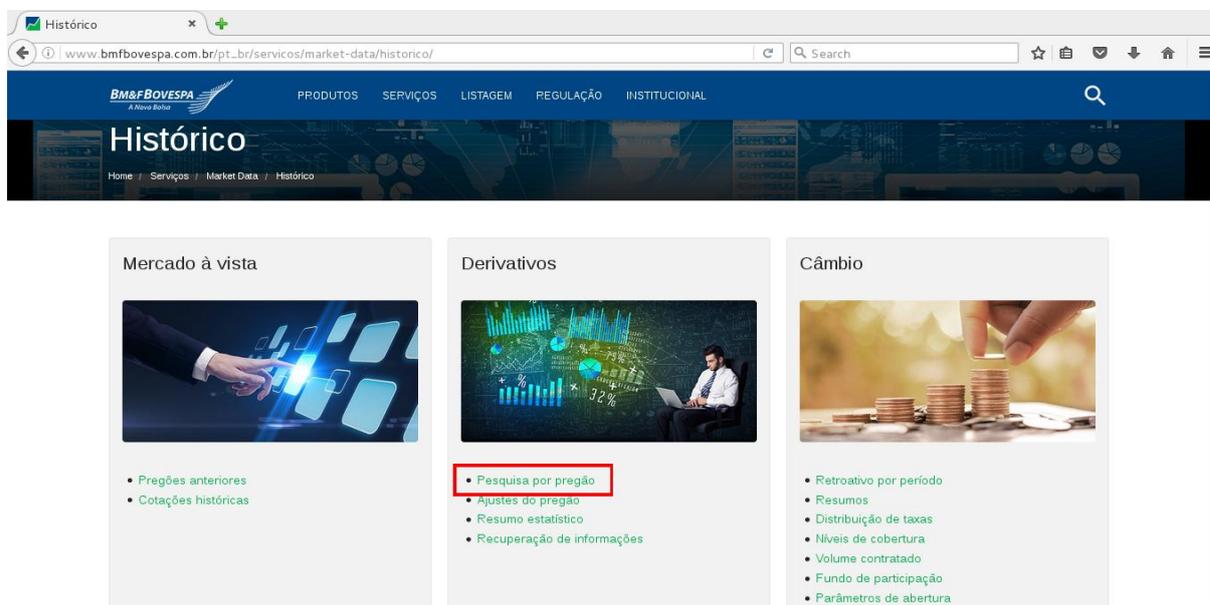


FIGURA 8 - Acesso área derivativos pesquisa por pregão.

Fonte: BM&FBovespa 2016.

Sendo encaminhado para uma página como na Figura 9 onde são exibidos diversos tópicos para pesquisa por pregão que produzem diferentes conteúdos.

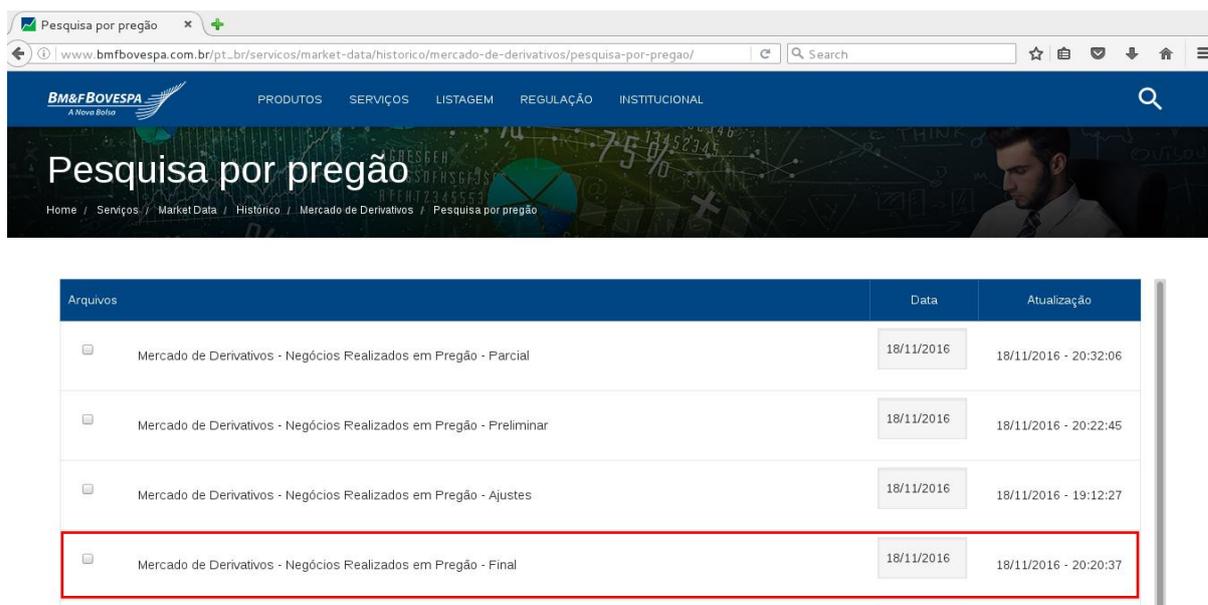


FIGURA 9 - Pesquisa por pregão.

Fonte: BM&FBovespa 2016.



Como pode ser observado na Figura 10 posteriormente não é possível distinguir, onde começa ou termina cada informação como data do pregão, código do contrato, tipo de registro e última cotação no dia. Totalizando 79 campos com informações que descrevem o que acontece durante o dia com um determinando contrato. Para dar a informação, a BM&FBovespa fornece um *layout* no formato de uma planilha como mostra a Figura 11.

	A	B	C	D	E	F	G
1		► MERCADO DE DERIVATIVOS					
2							
3		Negócios Realizados em Pregão					
4							
5		<b>Campos</b>	<b>Formato</b>	<b>Tamanho</b>	<b>Início</b>	<b>Fim</b>	<b>Observação</b>
6		Identificação da transação	N	6	1	6	Seqüência que ident.a transação
7		Complemento da transação	N	3	7	9	Fixo 001
8		Tipo de registro	N	2	10	11	Fixo 01
9		Data de geração do arquivo	N	8	12	19	AAAAAMDD
10		Tipo de negociação	A	2	20	21	PR (pregão)
11		Código da mercadoria	A	3	22	24	
12		Código do mercado	A	1	25	25	
13		Tipo da série	A	1	26	26	C (opç. compra) ou V (opç. venda)/* (fut)
14		Série (opç)/vencimento (fut)	A	4	27	30	
15		Hora de criação deste registro	N	6	31	36	HHMMSS
16		Data de vencimento (fut/opç)	N	8	37	44	AAAAAMDD

FIGURA 11 - Planilha descrição dos campos.

Fonte: Próprio autor.

A Figura 11 descreve os campos, formato da informação, tamanho, sequência de caracteres que demonstram onde começam e até onde eles vão . Por fim , a observação.

Para a mineração dos dados condensados no arquivo como mostra a Figura 10, foi escolhida a linguagem PHP (Hypertext Preprocessor), pelo motivo de domínio e conhecimento e simplicidade da mesma para elaborar um *script* que pode processar todos os arquivos usados, seguindo as instruções fornecidas pela BM&FBovespa observadas da Figura 11.

Com toda informação processada pelo *script* um total de 3246 arquivos, fez-se necessário armazenar todo este montante de dados em um Sistema de Gerenciamento de Banco de Dados (SGBD). Por domínio e conhecimento prévio foi decidida a utilização do SGBD MySQL para tal fim. Este, armazena aproximadamente 3 (2.933.436) milhões de registros. Entre eles estão todos os tipos de *commodities* brasileiras operadas diariamente. Com todos estes registros se faz necessário selecionar somente os contratos de café futuro e dentro dos contratos quais informações são de relevância para o estudo.

## 2.3 SELEÇÃO DOS ATRIBUTOS

Durante o estudo do trabalho foi estabelecido para uso das cotações diárias de contratos de café, o período de 01/01/2003 a 11/11/2016, somando um total de 717 semanas e 3.588 dias úteis de operações em bolsa. Para selecionar somente contratos referentes o café, observou que o campo código da mercadoria com tamanho de três caracteres e idêntico a ICF, foi filtrado um total de 97.210 registros.

Com o universo dos contratos de café selecionados se inicia o processo da seleção dos atributos ou campos que melhor expressam as informações no uso da RNA limitado nos seguintes campos, como a data de cotação do dia, o volume de negócios realizados na mesma data, valor em reais, o limite mínimo e máximo para as cotações do dia e o ajuste futuro.

Durante a seleção dos dados foi aplicado um filtro para resultar somente no conjunto de número racionais não negativos sem o zero ( $Q^{*+} = \{x \in Q \mid x > 0\}$ ). O resultado ocorreu em 10.896 cotações. Para o universo de treinamento foi utilizado 70% (7627) de todas as cotações e 30% (3269), para uso com validação da RNA, foram subdivididos em 30% de cada um dos 13 anos.

Com a escolha dos atributos para o treinamento da rede neural objetivando o treinamento, validação e previsão da cotação futura do encerramento diário, como veremos na sessão de treinamentos, foi realizado outro teste com objetivo em prever a oscilação diária, alterando ou aumentando os parâmetros de configuração para a rede MLP.

## 2.4 CONFIGURAÇÕES PARA REDE MLP

Para configuração e funcionamento da rede MLP se fez necessário estabelecer parâmetros iniciais como apresentados na Tabela 1, que descrevem as configurações e influenciam diretamente nos resultados da rede.

TABELA 1 - Parâmetros usados para configuração da MLP para realizar treinamento da rede e definir

<b>Atributo</b>	<b>Descrição</b>
setAutoBuild	Para conectar as camadas ocultas da rede define se como <i>true</i> . Definida como <i>false</i> para não decrementar a taxa de aprendizagem.
setNormalizeAttributes	Parâmetro para normalizar os atributos inseridos na rede para alcançar melhor desempenho.
setHiddenLayers	Determina a quantidade de camadas ocultas à rede terá.
setTrainingTime	Número de épocas para encerrar treinamento.
setLearningRate	Estabelece qual será a taxa de aprendizagem
setMomentum	Define como serão aplicados os pesos na rede

Fonte: Próprio Autor.

PETRY (2010) em seu trabalho comparou o uso de dois algoritmos de RNA da rede MLP com *Learning Vector Quantisation* (LRQ), para previsão de mercado de ações utilizando bandas de *Bollinge*, no qual obteve bons resultados, tendo assim adotado como ponto de partida para configurações iniciais das RNA a Tabela 2.

TABELA 2 - Configurações para RNA.

<b>Atributos</b>	<b>Valores</b>
Parâmetros para camadas ocultas (setHiddenLayers)	“a” = (atributos+classes)/2 “i” = atributos “t” = atributos + classes
Parâmetros para taxa de aprendizado (setLearningRate)	0,3 0,6 0,9 0,2
Parâmetros para ajuste de pesos (setMomentum)	0,23 0,27
Números de épocas (setTrainingTime)	1.000.000

Fonte: Próprio Autor.

Na Tabela 2 constam os atributos e os seus respectivos valores que são utilizados na sessão para treinamento e análise dos testes encontrados.

## 2.5 TREINAMENTOS E ANÁLISES DOS TESTES

Para estabelecer qual a rede que melhor obteve o resultados, foi necessário fazer inúmeros testes como o uso de três tipos de camadas ocultas distintas, três taxas de aprendizado diferentes, três ajustes em pesos incomuns, todas estas configurações iniciais realizando combinações entre si, variando também o número de épocas de treinamento em ordem progressiva começando por 1.000, 10.000, 100.000 e 1.000.000, no que gerou um resultante de 108 configurações e redes neurais, que dentre os melhores resultados detectados foram possíveis com 1.000.000 épocas de treinamento. A Tabela 3, reflete os melhores resultados atingidos nos treinamentos com uso das camadas ocultas.

TABELA 3 - Melhores RNA para camadas ocultas.

Nome para Rede	Valor p/ camada oculta	Erro absoluto	Erro quadrático	Erro Relativo Absoluto	Taxa Erro estimado
a	a	476,98	688,40	99,31	688,40
b	i	4526,46	4574,24	942,52	4574,24
c	t	4636,33	4682,99	965,39	4682,99

Fonte: Próprio Autor.

Nesta, estão os melhores resultados encontrados para cada tipo de camada oculta na RNA. O WEKA aceita informar valores para as camadas ocultas no formato numérico ou em letras que tem seus cálculos já pré-definidos. Camadas ocultas definidas que possuem o valor "a" como na rede "a" informada na primeira linha da Tabela 3, seguem a seguinte regra que são  $(\text{atributos} + \text{classes})/2$ . Esta é a quantidade de valores para entrada na rede, somados ao número de classes, onde por definição no WEKA é um conjunto para agrupar os atributos de entrada, dividido por dois resultando no valor de cada camada oculta.

Como no estudo foram utilizados seis valores para entrada, sendo eles data do pregão, dias de vencimento do contrato, volume em reais negociados, menor e

maior cotação negociada no dia junto ao valor da cotação diária. Para chegar ao valor usado na rede "a" aplica-se a regra definida somando as seis entradas com as classes que não foram empregadas no estudo, pois o agrupamento das informações podem afetar valores das redes optando assim pelo não uso, o resultado da soma é dividido por dois que resulta no valor de três camadas ocultas.

Na rede "b", utiliza-se o valor "i" para camadas ocultas que por definição do WEKA, é a mesma quantidade dos valores para entrada utilizada na rede "a" anteriormente citados sendo um total de 6 camadas ocultas.

O valor "t" para rede "c" é bem semelhante a regra usado para o atributo "a" na primeira rede com a diferença que a soma dos valores de entrada com a classe não é dividido por dois.

Entre os valores de camadas ocultas, o que melhor apresentou resultados foi o que demonstrou menor taxa de erro estimada foi a RNA "a" com uma taxa de 688,40.

Cada linha da Tabela 4 está relacionada um valor como o coeficiente de correlação que mede o grau de relação entre os valores das redes. O erro absoluto calculado pela RNA, o quadrático, o relativo absoluto e a taxa de erro estimada são fatores decisivos para escolha da rede.

TABELA 4 – Melhores RNA com valores para taxa de aprendizado.

<b>Nome para Rede</b>	<b>Valor p/ taxas de aprendizados</b>	<b>Erro absoluto</b>	<b>Erro quadrático</b>	<b>Erro Relativo Absoluto</b>	<b>Taxa Erro estimado</b>
a	0,3	669,33	872,35	139,37	872,35
b	0,6	476,98	688,40	99,31	688,40
c	0,9	491,52	660,88	102,34	660,88

Fonte: Próprio Autor.

A Tabela 4 mostra os resultados alcançados nos treinamentos das RNA para valores com as taxas de aprendizado. Para o critério de escolha da RNA foi usado o mesmo conceito aplicado para Tabela 3 escolhendo a rede com a menor taxa de erro estimada, a rede "c".

TABELA 5 - Melhores RNA com valores para ajustes de pesos.

Nome para Rede	Valor p/ ajuste de peso	Erro absoluto	Erro quadrático	Erro Relativo Absoluto	Taxa Erro estimado
a	0,20	491,52	660,88	102,34	660,88
b	0,23	497,39	662,41	103,56	662,41
c	0,27	476,98	688,40	99,31	688,40

Fonte: Próprio Autor.

Os resultados apresentados na Tabela 5 mostram os melhores valores para RNA com os ajustes de pesos que são aplicados no decorrer do treinamento da rede. Analisando entre elas a que teve menor taxa de erro estimada foi a rede “a”.

Os resultados para as RNA com a menor taxa de erro estimada para cada um dos parâmetros de configuração como camada oculta, taxa de aprendizado e ajuste de peso que foram apresentados anteriormente em cada uma das tabelas 3, 4 e 5, destaca-se respectivamente as redes “a”, “c” e “a” que podem ser vistas na Tabela 6.

TABELA 6 - Melhores RNA nos treinamentos.

Atributos	Nome p/ Rede	Valore configuração	Erro absoluto	Erro quadrático	Erro Relativo Absoluto	Taxa Erro estimado
Camadas Ocultas	a	a	476,98	688,40	99,31	688,40
Taxa de aprendizado	c	0,9	491,52	660,88	102,34	660,88
Ajustes de pesos	a	0,20	491,52	660,88	102,34	660,88

Fonte: Próprio Autor.

Na Tabela 6 nota-se que a rede “a” com atributo da camada oculta é a única que destaca das demais nos valores com erro absoluto, quadrático, relativo absoluto e taxa de erro estimada. Os menores valores selecionados para taxa de aprendizado e para o ajuste de pesos correspondem a uma mesma rede em que analisando com a perspectiva de cada um dos três principais atributos de configuração da rede com camadas ocultas, taxa de aprendizado e ajuste de pesos, sobre as 108 redes criadas.

Como podem ser observados nas tabelas posteriormente apresentadas as configurações individuais de cada parâmetro, destacou a melhor solução com uma taxa de erro estimada 660,88, camada oculta com atributo “a”, taxa de aprendizado em 0,9 e o ajuste de peso em 0,2. Encontrado assim o melhor aspecto para a RNA e possível avançar com as configurações definidas para à classificação e validação da rede.

## 2.6 CLASSIFICAÇÃO E VALIDAÇÃO DA REDE

Depois de realizado o treinamento e monitorado qual rede apresentou menor propensão da taxa de erros estimada, foi realizada a classificação dos 30% das cotações, os resultados podem ser observados na Figura 10.



FIGURA 12 - Treinamento e classificação pela RNA para cotações diárias.

Fonte: Próprio autor.

A Figura 12 retrata por meio das linhas em azul o valor real da cotação final de cada dia. Já a linha em vermelho equivale às cotações que foram classificadas pela RNA.

Para chegar a bons resultados obtidos com a classificação diária apresentada na Figura 12, inicialmente foi realizado treinamento de uma primeira RNA, seguindo

os mesmos métodos de configurações para a rede na classificação da cotação, contudo, foram utilizados mais parâmetros de entrada para a rede que são data do pregão, número de dias para vencimento do contrato, volume em reais operados no dia, menor e maior cotação negociada no dia, valor mínimo e máximo para cotações no dia e o valor de ajuste diário, com finalidade de classificar o ajuste diário, entre o período de 01/01/2007 a 11/11/2016, os resultados alcançados podem ser visualizados na Figura 13.

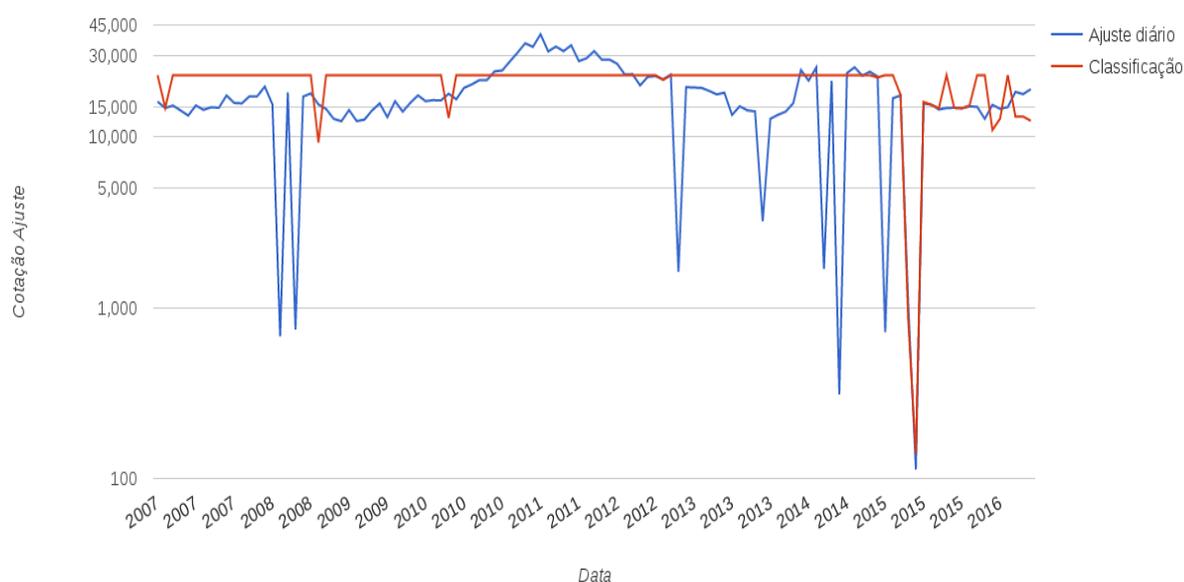


FIGURA 13 - Treinamento e classificação pela RNA para ajustes diários.

Fonte: Próprio autor.

Os ajustes diários reais são representados pela cor azul e os dados classificados pela RNA se encontram em vermelho. Observando a Figura 13 e analisando as duas linhas nota-se que a RNA tentou acompanhar as cotações reais dos ajustes diários, no então não foi satisfatória o treinamento da RNA para o ajuste diário, pois a formação deste valor é dada através da média ponderada dos últimos negócios efetuados nos minutos finais do fechamento do pregão, menos o valor do contrato. O ajuste diário é um importante cálculo que é divulgado pela BM&FBovespa após o fechamento do pregão, pois ,com base neste valor é que vai ser ajustado as diferenças entre os valores que vão ser creditados ou debitados nas contas de cada investidor seja um comprador ou vendedor.

Diante dos resultados encontrados com os processos de classificação e validação da rede com o ajuste diário em comparação com o resultado encorajador das cotações diárias, iniciou-se a etapa de classificação das previsões com a rede

que foi treinada com as cotações diárias. A rede treinada com os valores de ajustes diários não obteve bons resultados pelos treinamentos se mostrarem muito distantes do valor ideal.

## 2.7 CLASSIFICAÇÃO DAS PREVISÕES COM RNA

Após os resultados da classificação e validação da RNA deu-se início ao processo de previsão das cotações na qual foi estabelecido um período de previsão de cinco dias que vai do dia 24/10/2016 ao 28/10/2016, para os contratos com vencimento em dezembro de 2016. Para delimitar os valores que serão gerados aleatoriamente foram analisados os últimos 5 dias de cotações dos dias 17/10/2016 ao 21/10/2016. Para geração dos valores foram delimitados valores máximos e mínimos dispostos na Tabela 7.

TABELA 7 - Valores mínimos e máximo de 5 dias antes da previsão

<b>Menor Cotação</b>	<b>Maior Cotação</b>	<b>Volume em reais Negociados</b>	<b>Cotações</b>
1840	1872	R\$ 18.702.076,00	1867
1870	1898	R\$ 31.393.222,00	1883

Fonte: Próprio Autor.

Com os intervalos de cotações máximas e mínimas definidas, foram desenvolvidas funções para gerar as cotações aleatórias partindo do princípio que Petry (2010) e Rambo (2014) descrevem que a ocorrência das cotações passas tendem a se repetir em um futuro.

Os resultados obtidos para os próximos cinco dias podem ser observados pela Tabela 8 e Figura 9 com gráfico da previsão, onde estão às cotações reais geradas e classificadas pela RNA, na tabela além das cotações se destaca a margem de erro entre as cotações reais e previstas.

TABELA 8 - Previsão para os 24 a 28 de novembro de 2016.

Data	Cotação Real	Cotação Prevista		Margem de Erro
		Classificado pela rede	neural	
24/10/2016	1896	1895		0,05%
25/10/2016	1975	1897		3,94%
26/10/2016	1957	1892		3,30%
27/10/2016	1974	1895		3,96%
28/10/2016	1985	1893		4,62%

Fonte: Próprio autor.

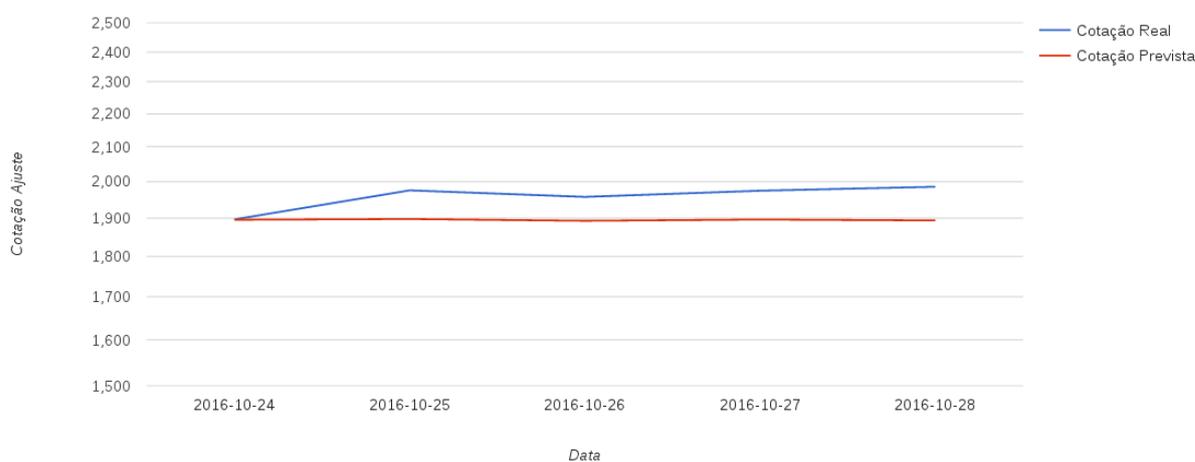


FIGURA 14 - Gráfico com a previsão para cinco próximos dias vezes a cotação real.

Fonte: Próprio autor

A Figura 12 apresenta graficamente os resultados apresentados na Tabela 8. A linha representada em vermelho demonstra as cotações previstas e a linha em azul o valor real da cotação nos dias, com uma ótima margem de aproximação entre as previsões e as cotações reais.

### 3 CONCLUSÕES

O presente trabalho propôs o prever a cotação diária de café na BM&FBovespa, utilizando cotações históricas para treinamento da RNA, gerando previsão que são classificações da RNA. Os resultados obtidos com o desenvolvimento foram satisfatórios, alcançando assim o objetivo principal do trabalho.

Para o desenvolvimento do trabalho foi utilizado a biblioteca WEKA. Por meio desta foi possível implementar o algoritmo MLP, criar uma RNA, treinar, testar, classificar e prever as cotações diárias futuras.

Para o desenvolver dos gráficos apresentados foi utilizado a *Application Programming Interface* (API) do Google chart<sup>1</sup>.

Em um mercado extremamente competitivo como o mercado de commodities, a representação de baixa ou alta de mercado, pode resultar em muitas perdas ou ganhos por parte de um investidor. Desta forma torna-se um fator importantíssimo a existência de ferramentas técnicas para auxiliá-lo em suas operações.

Com base neste argumento, percebeu-se a necessidade da empregabilidade da RNA. Frente ao treinamento e uso da mesma no decorrer do estudo, foi obtida uma margem média de erro de 1,77% para mais ou menos no universo selecionado. Já para as previsões das cotações houve uma margem média de erro em 3,17%, demonstrando assim que o algoritmo MLP com *backpropagation* são eficientes nas classificações das previsões, alcançando uma boa margem de erro. Dessa forma pode ser indicado como uma boa ferramenta no auxílio dos investidores na tomada de decisão em seus investimentos.

Para que se possa alcançar resultados ainda melhores que os obtidos podem ser alinhadas as cotações da *commodity* juntamente com outros índices como o Índice de *Commodities* Brasil (ICB) ou Índice Bovespa (Ibovespa), que podem ser fortes indicadores de cotações das *commodities*.

<sup>1</sup> Uma poderosa ferramenta para criação de gráficos, de simples compreensão desenvolvida gratuitamente pela Google (Google, 2016).

## TRABALHOS FUTUROS

Fica como sugestão para estudos em trabalhos futuros, a comparação entre RNAs com algoritmo MLP com *backpropagation*, frente a outros algoritmos na busca ao encontro de melhores resultados.

Pode ser realizada como estudo a identificação da ocorrência em repetições de séries temporais no mercado de commodities. Uma vez que os contratos possuem uma data não muito longas, podem sim haver repetições no mercado de ações como Guarnieri (2006) apresenta em seu estudo.

## REFERÊNCIAS

ABIC **Norma de Qualidade Recomendável e Boas Práticas de Fabricação de Cafés Torrados em Grão e Cafés Torrados e Moidos** Disponível em: <<http://www.abic.com.br/publique/media/Norma%20de%20qualidade.pdf>>.

Acesso em 28 de Ago. 2016.

BAPTISTA, Diana de Medeiros. **Integração e assimetrias na transmissão de preços de café arábica no Brasil**. (Doutorado) Tese apresentada para obtenção do título de Doutora em Ciências. Área de concentração: Economia Aplicada. Universidade de São Paulo Escola Superior de Agricultura “Luiz de Queiroz”, 2015.

BASILEIRO, Rodrigo de Carvalho. **SISTEMA AUTOMÁTICO PARA NEGOCIAÇÕES DE AÇÕES USANDO ALGORITMOS BASEADOS EM INTELIGÊNCIA COLETIVA E ANÁLISE TÉCNICA**. (Mestrado) Programa de Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, 2013.

BM&FBOVESPA. **Balanco de Operações**. Disponível em: <[http://www.bmfbovespa.com.br/pt\\_br/noticias/balanco-de-operacoes.htm](http://www.bmfbovespa.com.br/pt_br/noticias/balanco-de-operacoes.htm)> Acesso em 18 Ago 2015.

BM&FBOVESPA. **MANUAL DE PROCEDIMENTOS OPERACIONAIS DO SEGMENTO BOVESPA: AÇÕES, FUTUROS E DERIVATIVOS DE AÇÕES**. Disponível em: <<http://www.bmfbovespa.com.br/lumis/portal/file/fileDownload.jsp?fileId=8A828D294F270E45014FE2426FF239E7>> Acesso em 19 Ago 2016.

BOCCATO, Levy. **Novas Propostas e Aplicações de Redes Neurais com Estados de Eco**. (Mestrado) Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade de Campinas, CAMPINAS, 2013.

BORGES, Rafael Ribaski. **Plasticidade Sináptica em Redes Neurais**. (Tese) Graduação em Ciências, área de concentração Física, da Universidade Estadual de Ponta Grossa, PONTA GROSSA, 2016.

CANTO, Eduardo Leite do. **Ciências Naturais: aprendendo com o Cotidiano**. 3. ed. São Paulo: Moderna, 2009.

Conab. Séries históricas. Disponível em: <<http://www.conab.gov.br/conteudos.php?a=1252&>> Acesso em: 28 Novembro de 2016.

*Departmente of Conservation*. **Our overarching goal is to restore the abundance, distribution and genetic diversity of all kiwi species**. Disponível em: <<http://www.doc.govt.nz/nature/native-animals/birds/birds-a-z/kiwi/docs-work/>> Acesso em : 22 Novembro de 2016.

**DUARTE, Larissa Graziela Rauber. A valiação do emprego do café torrado como aromatizante na produção de cervejas.** (Mestrado) Pós Graduação da Escola de Engenharia de Lorena USP, LORENA, 2015.

**FRANCO, Marcelo Lemos Nunes. CLASSIFICAÇÃO DE NÓDULOS PULMONARES EM IMAGENS TOMOGRÁFICAS UTILIZANDO REDES NEURAIIS ARTIFICIAIS EM CASCATA.** (Mestrado) Engenharia Biomédica da Universidade Federal de Uberlândia, UBERLÂNDIA, 2015.

**GIACOMEL, Felipe dos Santos. Um Método Algorítmico para Operações na Bolsa de Valores Baseado em *Ensembles* de Redes Neurais para Modelar e Prever os Movimentos dos Mercados de Ações.** (Mestrado) Mestre em Ciência da Computação da UFRGS, PORTO ALEGRE, 2016.

**GUARNIERI, Odir Cantanhede. Um Estudo Empírico da Eficiência da Análise Técnica com Instrumento na Predição do Comportamento dos Preços das Ações: O Caso Embraer.** (Mestrado) em Gestão e Desenvolvimento Regional do Departamento de Economia, Contabilidade e Administração da Universidade de Taubaté, TAUBATÉ, 2006.

**JALE, Jader da silva. Commodities agrícolas do agronegócio brasileiro: análise multifractal e análise da complexidade diante da crise financeira mundial *subprime* 2008/2009.** (Doutorado) Programa de Pós-Graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco, RECIFE, 2015.

**KEMPER, Monika. Internacionalização do mercado de capitais brasileiro – a BM&FBOVESPA como centro financeiro regional.** (Tese) Departamento de Direito da PUC-Rio, RIO DE JANEIRO, 2011.

**KRIEGER, Paulo Eduardo. USO DE REDES NEURAIIS ARTIFICIAIS PARA PREDIÇÃO DA BOLSA DE VALORES.** (Bacharel) Em Ciência da Computação pela Universidade do Vale do Itajaí, ITAJAÍ, 2012.

**LEMES, LUCIENE ROSE. Modele de avaliação de risco para predição de preços de carne bovina utilizando inteligência computacional.** (Doutorado) Em Ciência na Área de Concentração em Ciência da Engenharia de Alimentos pela Universidade Federal de São Paulo, PIRASSUNUNGA, 2014.

**LUZ, Marcos Paulo Santos. ESTUDO DA RELAÇÃO DE FATORES CLIMÁTICOS COM A QUALIDADE DO CAFÉ NA MANTIQUEIRA DE MINAS.** (Mestrado) Pós-Graduação em Engenharia Agrícola pela Universidade Federal de Lavras, LAVRAS, 2014.

**MACIEL, Larissa Amanda Paiva. Müller, Carlos André da Silva. Silva, Rosália Maria Passos da. OPERAÇÃO EM MERCADOS FUTUROS E LUCRATIVIDADE DA PROPRIEDADE RURAL: ANÁLISE BASEADA NOS CUSTOS DE PRODUÇÃO DE CAFÉ.** ABCustos, São Leopoldo: Associação Brasileira de Custos, v. 11, n. 1, p.18-42, jan./abr. 2016.

MARTINS, Marco Antônio dos Santos Martins; METTE Frederike; MACEDO. Guilherme Ribeiro de. **AUTILIZAÇÃO DE REDES NEURAIIS ARTIFICIAIS PARA A ESTIMAÇÃO DOS PREÇOS DA PETROBRÁS PN NA BOVESPA**. ConTexto, Porto Alegre, v. 8, n. 14, 2º semestre 2008.

NETO, João Nunes de Mendonça. **Fractais e redes artificiais aplicados à previsão de retorno de ativos financeiros brasileiros**. (Mestrado) Universidade de São Paulo, 2014.

NETO, Luiz Gonzaga Lopes. **MODELAGEM DA DEGRADAÇÃO DE FENOL EM EFLUENTES AQUOSO ATRAVÉS DO PROCESSO DE FOTO-FENTON**. (Mestrado) Engenharia Química da Universidade Federal do Rio Grande do Norte, 2016.

NETO, Odilon José de Oliveira. **ANÁLISE DAS OPERAÇÕES DE HEDGER DO BOI GORDO NO MERCADO FUTURO DA MB&F PARA O ESTADO DE GOÍAS**. Salvador: Revista Gestão e Planejamento, v. 9, n. 1, p.77-93, jan./jun. 2008.

NICHI, Jaquiline. **Governança e Relato Integrado: discursos e práticas da Sustentabilidade corporativa**. (Mestrado) Escola de Artes, Ciência e Humanidades da Universidade de São Paulo, 2016.

NOGUEIRA, Irifrank Falção. **MINERAÇÃO DE DADOS NO BANCO DE DADOS DA COMISSÃO PRÓPRIA DE AVALIAÇÃO (CPA) DA FACULDADE DE BALSAS - UNIBALSAS**. (Bacharel) Em Sistemas da informação Faculdade de Balsas, BALSAS, 2013.

**O mercado de valores mobiliários brasileiro / Comissão de Valores Mobiliários**. 3. ed. Rio de Janeiro: Comissão de Valores Mobiliários, 2014.

PAGIATTO, Natália Ferrarezi. **ANÁLISE DE DITERPENOS E CAFEÍNA EM UMA COLEÇÃO DE ETIÓPIA DE COFFEA ARABICA**. (Mestrado) Programa de Pós-Graduação em Biotecnologia, Universidade Estadual de Londrina, LONDRINA, 2013.

PETRY, Adriano Cassaniga. **PROTÓTIPO PARA PREVISÃO DO MERCADO AÇÕES UTILIZANDO BANDAS DE BOLLINGER**. (Bacharelado) em Ciência da Computação pela Universidade Regional de Blumenau, BLUMENAU, 2010.

QUIRINO, Raphael de Souza Lima. **MERCADO DE CAPITAIS: A contabilidade como ferramenta análise para investimentos em bolsa de valores**. (Bacharelado) Graduação em Ciências Contábeis do Instituto de Pontifícia Universidade Católica de Minas Gerais, 2012.

RAMBO, Andrea Carneiro. **O PERFIL DO INVESTIDOR E MELHORES INVESTIMENTOS: da teoria à prática do mercado brasileiro**. (Bacharelado) em Ciências Econômicas pela Universidade Federal de Santa Catarina, FLORIANÓPOLIS, 2014.

ROCHA, Chrystian Teixeira. **Análise de integração de mercado dos principais países produtores de café.** (Mestrado) Programa de Pós-Graduação em Administração, Universidade Federal de Lavras, LAVRAS, 2009.

ROMEU, Marina Campanelli. **ANÁLISE DOS IMPACTOS DOS ESPECULADORES NOS RETORNOS DOS PREÇOS FUTUROS DAS PRINCIPAIS COMMODITIES AGRÍCOLAS EXPORTADAS PELO BRASIL.** (Mestrado) Escola Administração de Empresas de São Paulo, da Fundação Getúlio Vargas, 2014.

SERRA, Elpídio. **NOROESTE DO PARANÁ: o avanço das lavouras de cana e a nova dinâmica do uso do solo nas zonas de contrato arenito-basalto.** (Mestrado) Pós Graduação em Geografia pela Universidade Estadual de Maringá, 2010.

SILVA, Gabi Nunes. **REDES NEURAS ARTIFICIAIS: NOVO PARADIGMA PARA PREDIÇÃO DE VALORES GENÉTICOS.** (Mestrado) Pós Graduação em Estatística Aplicada e Biometria, Universidade Federal de Viçosa, Minas Gerais, 2014.

SILVA, Marcos Roberto Alves da. **DECISÕES DE ESTRUTURA DE CAPITAL NO BRASIL - UMA ABORDAGEM POR SETOR DE ATIVIDADE, FATORES ECONÔMICOS E DE MERCADO E DESEMPENHO EMPRESARIAL.**(Doutorado) Pós Graduação em Administração de Empresas da Universidade Presbiteriana Mackenzie, São Paulo, 2015.

SIMEÓN, Edgar Jhonny Amaya. **PROGNÓSTICO DE FALHAS BASEADO EM REDES NEURAS COM ESTADOS DE ECO.** (Doutorado) Em Sistemas mecatrônicos Universidade de Brasília Faculdade de Tecnologia Departamento de Engenharia Mecânica, Brasília 2015.

SOUZA, Igor Menezes Marinho de. **UM ESTUDO COMPARATIVO PARA PREVISÃO DA COTAÇÃO DE AÇÕES DA BM&FBOVESPA UTILIZANDO REDES NEURAS ARTIFICIAIS.** (Bacharel) Engenharia de Computação pela Escola Politécnica de Pernambuco, RECIFE, 2012.

STULP, Guilherme Toldo. **APLICAÇÃO E AVALIAÇÃO DE ALGORITMOS DE MINERAÇÃO DE DADOS PARA LOCALIZAR PADRÕES EM BANCOS DE DADOS.** (Bacharel) Em Ciência da Computação Universidade de Regional do Noroeste do estado do Rio Grande do Sul, IJUÍ, 2014.

TEIXEIRA, João Vitor Squillace. **Previsão de Índice da Bolsa de Valores do Estado de São Paulo utilizando Redes Neurais Artificiais.** (Bacharel) Ciência da Computação Universidade Federal de Lavras, LAVRAS, 2012.

TORRES, Nestor Paulo Palacios. **INTEGRAÇÃO ESPACIAL E EFICIÊNCIA NO MERCADO DE BOIGORDO.** (Mestrado) Economia da Universidade Federal de Santa Catarina, FLORIANÓPOLIS, 2005.

UNITED STATES DEPARTMENT OF AGRICULTURE – USDA. **Foreign Agriculture Service (FAS).** Production, Supply, and Distribution PS&D). Disponível em: <<http://apps.fas.usda.gov/psdonline/>> Acesso em: 12 ago. 2015.

VALE, Gustavo Bruno do. **RECONSTRUÇÃO E RECONHECIMENTO DE IMAGENS BINÁRIAS UTILIZANDO O ALGORITMO MÁQUINA DE *BOLTZMANN***. (Mestrado) Engenharia Elétrica da Universidade Federal de Uberlândia, Minas Gerais, 2016.

Z AidAN, Úrsula Ramos. **QUALIDADE DOS CAFÉS DA “REGIÃO DAS MATAS” EM FUNÇÃO DA VARIEDADE, DA ALTITUDE E DA ORIENTAÇÃO DA ENCOSTA DA MONTANHA**. (Mestrado) Pós-Graduação em Engenharia Agrícola pela Universidade Federal de Viçosa, VIÇOSA, 2015.