

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE CIÊNCIA DA COMPUTAÇÃO

JACKSON LUIS SCHOSSLER

**PREVISÃO DE VALOR DE UM ATIVO FINANCEIRO UTILIZANDO
REDES NEURAS ARTIFICIAIS PARA PROCESSAMENTO DE
DADOS DE SÉRIES TEMPORAIS**

TRABALHO DE CONCLUSÃO DE CURSO

MEDIANEIRA

2017

JACKSON LUIS SCHOSSLER

**PREVISÃO DE VALOR DE UM ATIVO FINANCEIRO UTILIZANDO
REDES NEURAS ARTIFICIAIS PARA PROCESSAMENTO DE
DADOS DE SÉRIES TEMPORAIS**

Trabalho de Conclusão de Curso apresentado ao Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Bacharel em Computação”.

Orientador: Prof. Me. Jorge Aikes Junior

Co-orientador: Prof. Dr. Arnaldo Candido Junior

MEDIANEIRA

2017



TERMO DE APROVAÇÃO

**PREVISÃO DE VALOR DE UM ATIVO FINANCEIRO UTILIZANDO REDES
NEURAS ARTIFICIAIS PARA PROCESSAMENTO DE DADOS DE SÉRIES
TEMPORAIS**

Por

JACKSON LUIS SCHOSSLER

Este Trabalho de Conclusão de Curso foi apresentado às 13:50h do dia 13 de Junho de 2017 como requisito parcial para a obtenção do título de Bacharel no Curso de Ciência da Computação, da Universidade Tecnológica Federal do Paraná, Câmpus Medianeira. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Jorge Aikes Junior
UTFPR - Câmpus Medianeira

Prof. Arnaldo Candido Junior
UTFPR - Câmpus Medianeira

Prof. César Angonese
UTFPR - Câmpus Medianeira

Prof. Paulo Lopes de Menezes
UTFPR - Câmpus Medianeira

A folha de aprovação assinada encontra-se na Coordenação do Curso.

RESUMO

SCHOSSLER, Jackson Luis. PREVISÃO DE VALOR DE UM ATIVO FINANCEIRO UTILIZANDO REDES NEURAIAS ARTIFICIAIS PARA PROCESSAMENTO DE DADOS DE SÉRIES TEMPORAIS. 57f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná, Medianeira, 2017.

A previsão de Séries Temporais representa um grande papel na economia, em especial, no ramo de investimento baseado na compra e venda de ativos de empresas com o capital aberto, onde e de suma importância prever se o valor dos ativos irá subir ou baixar, afim de maximizar os lucros ou minimizar as perdas. Uma série temporal é um conjunto de observações sobre um tipo de dado, ordenado no tempo e registrado em períodos regulares. A suposição básica que da rumo a análise de séries temporais é que há um sistema relacionado com o tempo, que exerceu influência sobre os dados no passado e possivelmente continuará exercendo no futuro, dando origem a padrões não aleatórios, podendo ou não ser interpretado por gráficos, processos estatísticos e matemáticos ou redes neurais artificiais. Os modelos mais comuns para a predição do mercado financeiro são matemáticos ou estatísticos e dependem de um conhecimento técnico avançado e muitas adaptações de modelo.

Palavras-Chaves: Séries temporais, previsão, mercado.

Abstract

SCHOSSLER, Jackson Luis. 57f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná, Medianeira, 2017.

The prediction of the Time Series(ST) plays a major role in the economy, especially in the branch of investment based on the purchase and sale of stocks of publicly traded companies, where it is very important to predict if the stock price will rise or fall, in order to maximize profits or minimize losses. A time series is a set of observations on a data type, ordered in time and recorded in regular periods. The basic assumption leading to time-series analysis is that there is a time-related system that exerted influence on data in the past and possibly will continue to exercise in the future, giving rise to non-random patterns, which may or may not be interpreted by graphs , Statistical and mathematical processes or artificial neural networks. The most common models for predicting the financial market are mathematical or statistical and depend on advanced technical knowledge and many model adaptations.

Keywords: Time series, prediction, market

AGRADECIMENTOS

A minha família pelo apoio e incentivo incondicional. Aos Professores e colegas, que se tornaram grandes amigos nesta jornada. E a todos que direta ou indiretamente fizeram parte desta fase, o meu muito obrigado.

LISTA DE FIGURAS

FIGURA 1	– Fluxo entre famílias e empresas.	13
FIGURA 2	– Garrafas de vinho vendidas na Austrália.	20
FIGURA 3	– Modelo de previsão	22
FIGURA 4	– Sazonalidades e tendências.	24
FIGURA 5	– Processo de descoberta de conhecimento em banco de dados	25
FIGURA 6	– Neurônio Biológico.	35
FIGURA 7	– Google Finance	40
FIGURA 8	– Gráfico das variações previstas e reais	52

LISTA DE TABELAS

TABELA 1	– Quadro de comparação entre técnicas de previsão.	23
TABELA 2	– Tabela de dados divididos em 5 minutos.	46
TABELA 3	– Tabela dos resultados do grupo com o valor de fechamento.	47
TABELA 4	– Tabela dos resultados do grupo com o valor de todos atributos.	48
TABELA 5	– Tabela dos resultados dos testes para o dia 15/02/2017.	50
TABELA 6	– Tabela dos resultados dos testes para o dia 15/02/2017 e cálculos.	51
TABELA 7	– Tabela dos resultados e lucro no dia 15/02/2017.	52

LISTA DE SIGLAS

AI	Artificial Intelligence
API	Application Programming Interface
CSV	Comma Separated Values
GPL	General Public License
GUI	Graphical User Interface
IA	Inteligencia Artificial
KDD	Knowledge Discovery in Databases
ON	Ações Ordinárias
PN	Ações Preferenciais
RNA	Redes Neurais Artificiais
ST	Séries Temporais
VLSI	Very Large Scale Integration
WEKA	Waikato Environment for Knowledge Analysis

SUMÁRIO

1	INTRODUÇÃO	10
1.1	OBJETIVOS GERAL E ESPECÍFICOS	11
1.1.1	Objetivos Específicos	11
2	MERCADO FINANCEIRO	12
2.1	MERCADO DE CAPITAIS	14
2.2	AÇÕES	16
2.3	NASDAQ	17
2.3.1	Segmentos de negócios	17
2.3.2	Considerações Finais	18
3	SÉRIES TEMPORAIS	19
3.1	PREVISÃO DE SÉRIES TEMPORAIS	21
3.2	DESCOBERTA DE CONHECIMENTO	24
3.2.1	Coleta de dados	25
3.2.2	Pré-Processamento	26
3.2.3	Mineração de dados temporais	27
3.2.4	Transformação de Dados	28
3.2.5	Avaliação e interpretação	29
4	INTELIGENCIA ARTIFICIAL	30
4.1	ABORDAGENS NO ESTUDO DE INTELIGÊNCIA ARTIFICIAL	32
4.1.1	Agindo de forma humana	32
4.1.2	Pensando de forma humana	33
4.1.3	Pensando racionalmente	33
4.1.4	Agindo racionalmente	34
4.2	REDES NEURAS	34
4.2.1	Perceptron	37
4.2.2	Multilayer Perceptron	37
4.3	CONSIDERAÇÕES FINAIS	38
5	MATERIAIS E MÉTODOS	39
5.1	COLETA DE DADOS	39
5.2	FORMATO DOS DADOS	40
5.3	AMBIENTE	41
5.4	WEKA	41
5.4.1	Interface do usuário	42
5.4.2	MultiLayer Perceptron	43
5.4.3	Parametros de Rede Neural Artificial no Weka	43
5.5	DIVISÃO DAS SÉRIES TEMPORAIS	45
5.6	CONSIDERAÇÕES FINAIS	46
6	RESULTADOS OBTIDOS E DISCUSSÃO	47
6.1	TESTES COM DADOS REAIS	49
7	CONCLUSÃO	53
7.1	LIMITAÇÕES	54

7.2 TRABALHOS FUTUROS	54
REFERÊNCIAS	56

1 INTRODUÇÃO

A previsão de Séries Temporais (ST) representa um grande papel na economia, em especial, no ramo de investimento baseado na compra e venda de ativos de empresas com o capital aberto, onde é de suma importância prever se o valor dos ativos irá subir ou baixar, afim de maximizar os lucros ou minimizar as perdas. Uma série temporal é um conjunto de observações sobre um tipo de dado, ordenado no tempo e registrado em períodos regulares. A suposição básica que dá rumo a análise de séries temporais é que há um sistema relacionado com o tempo, que exerceu influência sobre os dados no passado e possivelmente continuará exercendo no futuro, dando origem a padrões não aleatórios, podendo ou não ser interpretado por gráficos, processos estatísticos e matemáticos ou redes neurais artificiais (RNA). Os modelos mais comuns para a predição do mercado financeiro são matemáticos ou estatísticos e dependem de um conhecimento técnico avançado e muitas adaptações de modelo (REIS, 2007).

A ideia de previsão do mercado de ações não é nova, investidores bem sucedidos buscam antecipar o mercado interpretando parâmetros externos, como opinião pública, indicadores econômicos e clima político, mas nem todos os fatores que influenciam no valor das empresas podem ser detectados facilmente. Muitas vezes tendências dos dados podem prever possíveis oscilações, um modelo que interprete esses dados corretamente com o mínimo de erro possível, pode se tornar extremamente útil na hora de comprar ou vender ações no mercado financeiro.

Empresas do ramo de tecnologia estão ganhando cada dia mais espaço e força no mercado financeiro, algumas delas, inclusive, são detentoras dos maiores capitais do mercado global. A NASDAQ Stock Market é um mercado de ações automatizado onde estão listadas mais de 2800 ações de diferentes empresas. A NASDAQ se caracteriza por reunir empresas de alta tecnologia em eletrônica, informática, telecomunicações e biotecnologia.

O interesse da inteligência artificial (IA) no mercado financeiro vem crescendo, justamente a fim de se beneficiar das possíveis previsões de oscilações no valor das ações, e com a utilização de RNA, têm apresentado bons resultados quando o objetivo é prever e classificar dados. De acordo com o ajuste do modelo e da RNA utilizada, elas podem extrair as características mais importantes de uma determinada ST.

Uma RNA é um paradigma de processamento de informação que é inspirado na forma como os sistemas biológicos nervosos funcionam, é composto por uma grande quantidade de elementos de processamento altamente interligados (neurônios) que trabalham em conjunto para resolver problemas específicos através do processo de aprendizagem. Aprendizagem em sistemas biológicos envolve adaptações das conexões sinápticas que existem entre os neurônios interligados, o que também ocorre nos neurônios artificiais (KOVÁCS, 1996).

1.1 OBJETIVOS GERAL E ESPECÍFICOS

O objetivo geral deste trabalho consiste em fazer uso de redes neurais artificiais em séries temporais para previsão do valor de um ativo na bolsa de valores para o auxílio às tomadas de decisões dos investidores para compras e vendas a curto prazo. Para se alcançar este objetivo geral, foram definidos os seguintes objetivos específicos.

1.1.1 Objetivos Específicos

- Obter dados históricos do valor de um ativo financeiro de uma determinada empresa e utiliza-lo como parâmetro para predições;
- Modelar o sistema;
- Selecionar um software que efetue o processamento dos dados;
- Experimentar, avaliar e modelar as melhores configurações para as tarefas pretendidas;
- Analisar a precisão dos resultados obtidos da RNA.

2 MERCADO FINANCEIRO

Em qualquer sociedade econômica moderna, o sistema financeiro é parte integrante e de extrema importância. O desejo de consumo de uma sociedade é ilimitado, porém os recursos necessários para produzir estes bens de consumo são limitados. O entendimento sobre a alocação destes recursos é importante pois a oferta quase sempre é menor do que a demanda. É fundamental ter noções básicas sobre o funcionamento da economia, para depois entender o sistema financeiro e o funcionamento dos mercados (MOBILIÁRIOS, 2014).

Atualmente, as empresas estão inseridas em um mercado altamente competitivo, devido o acirramento da competitividade global o cenário econômico mundial tem passado por transformações significativas, internacionalização dos mercados de capitais e inovações financeiras. Assim, as empresas devem desenvolver estratégias tendo como objetivo manter sua sobrevivência neste mercado (LEMOS et al., 2013).

Os consumidores estão se tornando cada vez mais exigentes em relação aos produtos, ao buscarem sempre melhor qualidade com o menor impacto negativo no meio ambiente durante o processo produtivo, logo o avanço da tecnologia da informação e a globalização derrubam as barreiras geográficas possibilitando a disseminação desta mudança de comportamento social (LEMOS et al., 2013).

Este novo paradigma teve influência sobre o papel das empresas no âmbito social, tendo em vista que a sociedade também é uma parte interessada à empresa. Observam-se duas visões relativas ao papel das companhias: a clássica que restringe como função das empresas a maximização do lucro para os acionistas; e a socioeconômica, que relaciona o bem-estar social ao papel das entidades, não atendo-se somente no lucro (LEMOS et al., 2013).

Os três pilares da economia são empresas, famílias e o governo. As famílias são parte integrante da mão de obra, em troca do salário. As empresas, de modo geral, utilizam desta mão de obra para produzir bens de consumo, que serão vendidos novamente para estas famílias, fazendo o capital girar. Na outra mão, tem o governo que recolhe impostos, e os transformam em projetos sociais e serviços básicos (MOBILIÁRIOS, 2014).

Na Figura 1 pode ser visualizado o fluxo entre famílias e empresas, onde as famílias geram receita das empresas à partir das despesas de consumo, em seguida recebem novamente

pelos serviços prestados (salário).

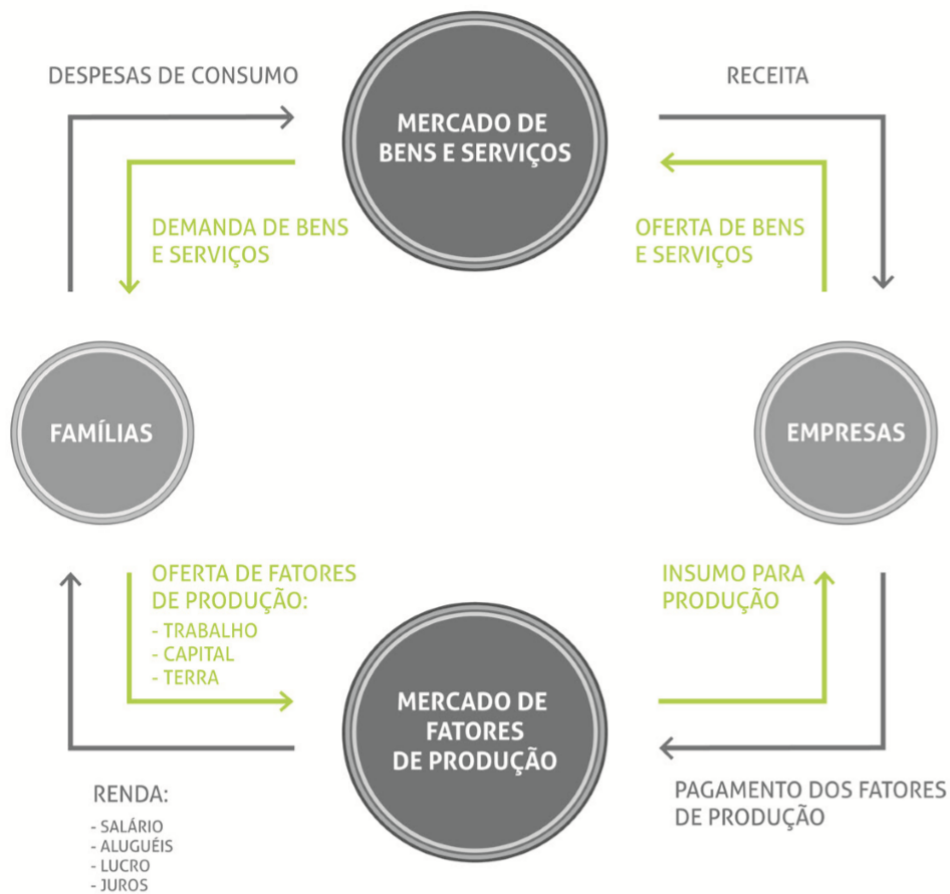


Figura 1 – Fluxo entre famílias e empresas.

Fonte: (MOBILIÁRIOS, 2014)

Além da troca da mão de obra por salários e salários por bens de consumo, o sistema financeiro também funciona com poupança e investimento. Por exemplo, uma família pode gastar menos do que ganha, seja por segurança financeira, aposentadoria, ou uma futura compra de bens, gerando assim uma poupança. Porém, esta mesma família, em determinado momento, pode estar gastando mais do que ganha, sendo necessário a tomada de empréstimos para quitar suas despesas. Isso também pode ser aplicado com as empresas, que podem guardar dinheiro para ter um capital de giro sempre disponível, porém pode precisar de mais dinheiro do que ganha, para investir em máquinas e equipamentos. Não diferente das famílias e empresas, o governo pode poupar um capital, para segurança financeira, mas também pode ter gastos maiores que sua renda, precisando tomar dinheiro emprestado de agentes financeiros internos ou externos, gerando uma dívida pública interna ou externa (MOBILIÁRIOS, 2014).

Para intermediar essas operações, surgiram instituições especializadas. Essas instituições pegam emprestado o dinheiro daqueles que poupam, pagando uma certa quantia

representada pelos juros e emprestam para quem precisa, trivialmente a uma taxa maior, a fim de lucrar com essa diferença. Com a formação destas instituições, foram se desenvolvendo novos instrumentos, sistemas e produtos, para controlar e desenvolver todo este mercado.

Uma das definições de Sistema Financeiro é o conjunto de instituições e instrumentos que viabilizam o fluxo financeiro entre os poupadores e os tomadores na economia. Fazem parte deste sistema todas as instituições e instrumentos que de alguma maneira, trabalham com o fluxo de empréstimos. É fácil compreender a importância deste sistema para o funcionamento adequado e crescimento econômico de uma nação (MOBILIÁRIOS, 2014).

2.1 MERCADO DE CAPITAIS

O mercado de capitais é uma alternativa vantajosa de investimento e financiamento de empresas para o crescimento patrimonial das unidades econômicas. No ramo do mercado financeiro, os investidores, ao emprestarem recursos diretamente para empresas, adquirem títulos que estabelecem as condições do negócio, chamados valores mobiliários. Existem três tipos de títulos, os de dívida, que se estabelece uma relação de crédito entre as entidades poupadora e tomadora, podem ser títulos patrimoniais, ou de capital, no qual os investidores se tornam sócios do negócio, como é o caso do investimento através da compra de ações, sendo que tanto a parte poupadora (compra as ações) como a tomadora (vende suas ações) tem direitos e deveres inerentes. Os valores mobiliários podem ser adquiridos pelos investidores direta ou indiretamente por meio de investimento coletivo, como clubes de investimento ou fundos (MOBILIÁRIOS, 2014).

Segundo EBC (2015) a bolsa de valores gerencia as compras e vendas de ações ou títulos que equivale a pequenas fatias do capital de uma determinada empresa. Quando um investidor compra uma ação, se torna de certa forma um sócio da empresa que comprou esta ação, podendo lucrar ou perder dinheiro de acordo com a oscilação de valor desta empresa na bolsa de valores. Em alguns casos quem compra as ações pode influenciar até em decisões da diretoria, dependendo da porcentagem que possui de ações de uma empresa.

A responsabilidade pelo pagamento dos juros obtidos é da emissora. As instituições financeiras não assumem nenhuma responsabilidade pelo cumprimento ou não das obrigações estabelecidas, pois o processo de colocação dos títulos no mercado é realizado pelas emissoras (MOBILIÁRIOS, 2014).

Da ótica dos investidores, o mercado de capitais surge como alternativa às tradicionais aplicações em produtos oferecidos pelos bancos ou governo. É nesse tipo de mercado que os investidores podem decidir em quantos e quais empreendimentos devem aplicar seu dinheiro, desde que assumam os riscos decorrentes. O risco é um fator de extrema relevância na hora de investir, quanto maior a diversidade dos investimentos, menor o risco de se ter grandes perdas. Nos títulos patrimoniais se tem maior rentabilidade, porém os riscos também são maiores, pois é o investidor que opera os recursos aplicados (O'NEIL, 2009).

Uma definição do mercado de capitais consiste que ele é um segmento do mercado financeiro o qual são criadas condições para que empresas captem recursos diretamente de um poupador ou investidor, por meio da emissão de instrumentos financeiros, a fim de financiar suas atividades e viabilizar projetos de investimentos. Segundo MOBILIÁRIOS (2014) “O mercado de capitais tem uma grande importância no desenvolvimento do país, pois estimula a poupança e o investimento produtivo, o que é essencial para o crescimento de qualquer sociedade econômica moderna.”

Uma empresa pode ter ações na bolsa de valores, assim é chamada de Sociedade Anônima (capital aberto), ao contrário, por exemplo, de uma empresa limitada (Ltda), cujo capital não pode ser negociado em bolsas de valores (capital fechado). Geralmente empresas de grande porte tem seu capital aberto e dependem de muitos investidores, já empresas de pequeno e médio porte, na maioria das vezes, têm seu capital fechado (INFOMONEY, 2016).

Através do lançamento de ações no mercado, uma empresa pode aumentar substancialmente o número de acionistas (sócios), tendo a vantagem de obter mais recursos para possíveis investimentos, que provavelmente não estariam disponíveis caso esta empresa fosse de capital fechado. Do ponto de vista de quem compra estas ações, há a possibilidade de lucrar com a compra e venda dos ativos, caso a empresa valorize (INFOMONEY, 2016).

Ter noção das principais características dos ativos financeiros em um mercado de capitais, facilita o investidor a planejar a alocação de seus investimentos onde melhor se ajustem aos seus objetivos, planejamento de riscos e prazos. No mercado de capitais, os ativos podem ser chamados de valores mobiliários, que podem ser qualquer contrato de investimento coletivo, que dá direito de participação, parceria ou remuneração, dos quais os rendimentos são resultados do esforço do empreendedor ou de terceiros (MOBILIÁRIOS, 2014).

2.2 AÇÕES

No mercado financeiro, as ações diferenciam basicamente pelo direito concedido a cada um de seus acionistas. Elas são divididas em dois grandes grupos: Ações ordinárias (ON) e as ações preferenciais (PN). Os dois tipos de ações são também nominativas, ou seja, cada detentor é listado nos livros de registro de uma empresa (TUBINO, 2006). Segundo Infomoney (2016), os dois tipos de ações, preferenciais (PN) e ordinárias (ON), podem ser definidas como:

Ações preferenciais: São chamadas assim porque dão preferência aos acionistas para o pagamento de dividendos ou no caso de liquidação da empresa. Isto significa que caso a empresa declare falência, ou algum outro evento que acarrete na liquidação da empresa, os que detêm ações preferenciais têm maiores chances de recuperar parte do valor investido do que os donos das ações ordinárias. Geralmente são as ações preferenciais que têm maior liquidez, porque permitem à empresa emitir estas ações sem ser necessário conceder direito a voto, assim os donos não precisam correr o risco de perderem o controle da empresa (TUBINO, 2006; INFOMONEY, 2016).

Ações ordinárias: São os tipos de ações que dão direito a voto ao acionista. Algumas vezes, o acionista também tem direito a veto, podendo contestar algumas ações dos sócios que estão no controle da empresa. Também é importante ressaltar que apesar de um acionista possuir direito a voto, ele não é responsável pelas dívidas da empresa (algo que é comum quando a empresa é de capital fechado). Desta forma, o pior cenário possível é que o acionista perca todo o dinheiro investido, caso aconteça algo com a empresa.

Além destas duas categorias de ações, as empresas podem emitir algumas outras classes de ações. Por exemplo, analisando as ações listadas na BOVESPA, não são listadas somente as ações ordinárias e ações preferenciais. Isso ocorre porque além das distinções básicas ON e PN, as empresas podem diferenciar as classes das suas ações de acordo com o critério de restrição quanto à posse de ações, de distribuição de dividendos, entre outros (INFOMONEY, 2016).

2.3 NASDAQ

A Nasdaq Stock Market é um mercado de ações automatizado onde estão listadas mais de 2800 ações de diferentes empresas. A empresa teve início em 1971 como uma pequena trocadora de ações com sede nos Estados Unidos, e hoje é reconhecida em todo o mundo como um provedor de serviços e tecnologia financeira, comercial e de informações para os mercados de capitais, com mais de 50 escritórios em 26 países. Em 2007 a NASDAQ se funde com a OMX, operadora de câmbio líder nos países nórdicos, para diversificar e expandir sua presença global, e complementa seu nome para NASDAQ OMX Group (NASDAQ, 2016).

Com base na experiência de mercado de capitais, conhecimento tecnológico, a NASDAQ tem aumentado e diversificado os seus produtos e serviços. Dentre os serviços oferecidos, estão (NASDAQ, 2016) :

- Desenvolver tecnologias eficientes e confiáveis que orquestram e facilitam a atividade nos mercados de capitais;
- Gerenciar as complexidades e custos de negocio em escala global;
- Fornecer dados e ferramentas que impulsionam a tomada de decisão na hora de investir no mercado de capitais.

O principal ramo das empresas listadas neste mercado de ações são de tecnologia em eletrônica, informática, telecomunicações e biotecnologia (NASDAQ, 2016).

2.3.1 Segmentos de negócios

O modelo da NASDAQ é construído estrategicamente em torno de quatro segmentos de negócios que gerem todas suas ações e serve como base para todo serviço realizado pela NASDAQ, esses quatro segmentos de negócios são: Tech, Trade, Intel, List (NASDAQ, 2016).

Tech: As soluções tecnológicas representa o software comercial. Ele oferece tecnologia de troca que suporta a industria financeira em todos os aspectos de mercado. O negócio de mercado de tecnologia é a maior provedora do mundo de soluções tecnológicas de câmbio e um fornecedor superior de soluções de corretagem. A cada dez transações de títulos no mundo, uma é utilizada em alguma solução de software comercial;

Trade: O segmento de comércio é construído através de compradores e vendedores juntos

em um ambiente aberto e transparente que representa uma das maiores redes mundiais de trocas, operando em 26 mercados em todo mundo. Rápido e confiável, têm grande capacidade de roteamento e excelente serviço ao cliente;

List: O modelo de lista revolucionário que a NASDAQ vem desenvolvendo há mais de 40 anos serve como padrão para os mercados mundiais. Trazem uma seleção diversificada de mais de 3700 empresas, com US\$ 9,3 trilhões em valor de mercado, representando indústrias como varejo, saúde, finanças e tecnologia;

Intel: Com os dados de mercado do segmento de Intel, os clientes em todo o mundo têm a visão e os recursos necessários para um melhor entendimento sobre transparência de mercado e automaticamente compreendem melhor os investimentos e decisões a serem tomadas.

Estes quatro segmentos de negócios são fundamentais para a gestão dos serviços, pois eles servem como base para o investidor entender como funciona o negócio que pretende investir e o auxilia na tomada de decisão.

2.3.2 Considerações Finais

Todos estes segmentos em que a NASDAQ opera seus negócios, fazem dela a maior bolsa de valores no setor de tecnologia, as ferramentas e técnicas que ela oferece, auxiliam para que o investidor tenha mais segurança e confiança na hora de aplicar seus fundos. No próximo capítulo será apresentado o que são séries temporais, como funcionam e como podem ser aplicadas no mercado de ações.

3 SÉRIES TEMPORAIS

Uma série temporal (ST) é um conjunto de observações sobre um tipo de dado, ordenado no tempo e registrado em períodos regulares. A suposição básica que dá rumo a análise de séries temporais é que há um sistema relacionado com o tempo, que exerceu influência sobre os dados no passado e possivelmente continuará exercendo no futuro, dando origem a padrões não aleatórios, podendo ou não ser interpretado por gráficos, processos estatísticos e matemáticos ou redes neurais artificiais (REIS, 2007). A previsão de séries temporais desempenha um grande papel na economia, em especial, no ramo de investimento baseado na compra e venda de ativos de empresas com o capital aberto, onde é de suma importância prever se o valor dos ativos irá subir ou baixar, afim de maximizar os lucros ou minimizar as perdas. Os modelos mais comuns para a predição do mercado financeiro são matemáticos ou estatísticos e dependem de um conhecimento técnico avançado e muitas adaptações de modelo (REIS, 2007).

Segundo Morettin (1981), são exemplos de séries temporais:

1. Estimativas trimestrais do Produto Nacional Bruto;
2. Valores diários de temperatura da cidade de São Paulo;
3. Índices diários da bolsa de valores;
4. Quantidade mensal de chuva na cidade de Avaré;
5. Valores mensais de venda de automóveis no Brasil;
6. Um registro de marés no porto de Santos.

Os cinco primeiros exemplos apresentam séries temporais discretas, já o sexto exemplo é uma série temporal contínua. Geralmente, uma série temporal discreta é obtida por uma amostragem de uma série temporal contínua mas em intervalos de tempos iguais (delta). Assim, para analisar a série do exemplo seis, faz-se necessário amostrá-la em intervalos de tempos iguais (de hora em hora por exemplo).

Basicamente, há dois enfoques usados na análise de séries temporais. Nos dois casos, o objetivo é construir modelos para as séries, com um intuito determinado. No primeiro enfoque, diz-se que a análise é feita no domínio temporal e os modelos são paramétricos, ou seja, com número finito de parâmetros. No segundo enfoque, analisa-se o domínio de frequências e os

modelos propostos são modelos não paramétricos (MORETTIN, 1981).

No domínio de frequências, tem-se a análise espectral, que por sua vez, é muito aplicado em ciências físicas e engenharia, e basicamente consiste em decompor uma série dada em componentes de frequência e onde a existência de um espectro é a característica fundamental (MORETTIN, 1981).

Para melhor compreensão dos fundamentos, pode-se aplicar o exemplo a seguir. Parte-se da premissa que se deseja contabilizar quantas garrafas de vinho tinto foram vendidas mensalmente na Austrália durante o período de 1980 e 1992. Pode-se ter uma noção visual à partir do gráfico de garrafas de vinho vendidas na Austrália.

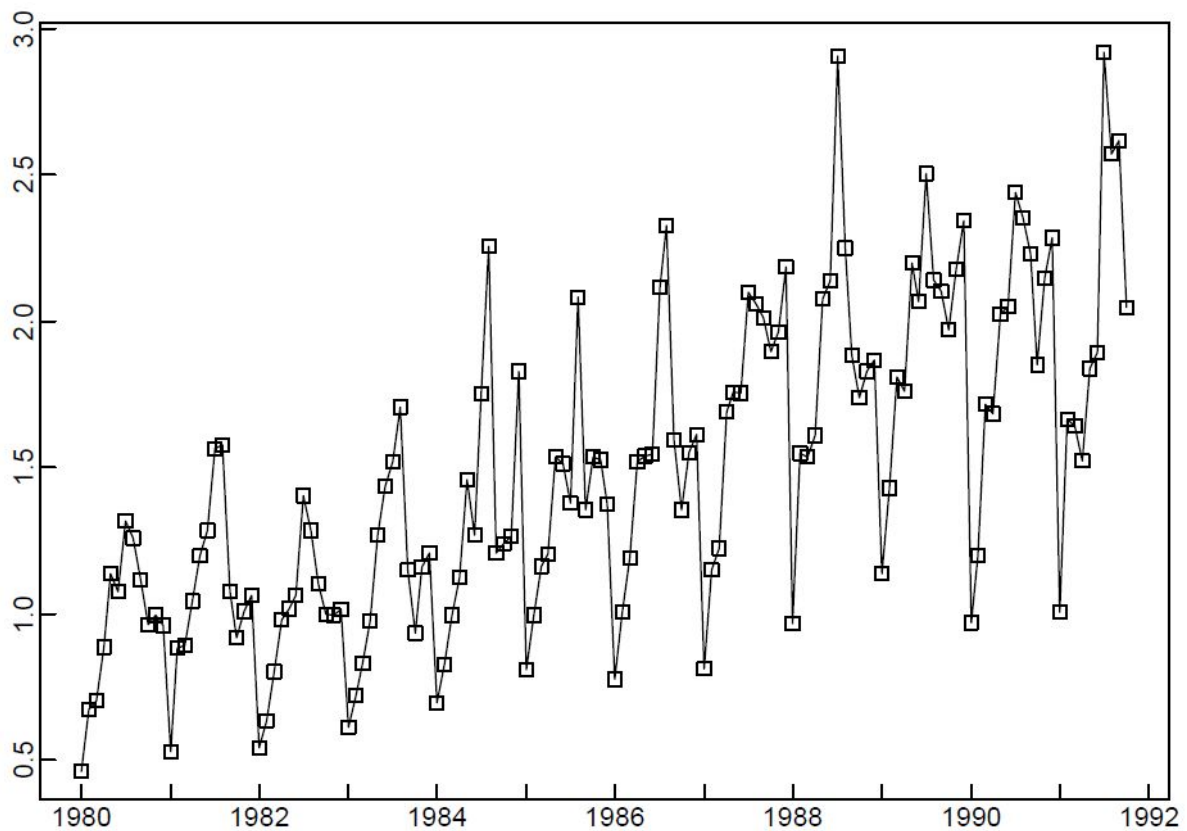


Figura 2 – Garrafas de vinho vendidas na Austrália.

Fonte: (MORETTIN, 1981)

Após se ter uma noção do que é uma série temporal e ter visto alguns exemplos, pode-se pensar em modelos hipotéticos de probabilidade para representar os dados. Depois de ter escolhido um modelo, é também possível estimar parâmetros para este modelo. O objetivo desses modelos é tentar identificar os componentes de uma série e como ela é constituída estruturalmente para tentar realizar alguma tarefa de interesse. Uma vez que este modelo é satisfatório, existe uma imensidão de áreas e maneiras em que este modelo pode ser trabalhado,

dependendo do campo da aplicação. O modelo pode simplesmente servir para fornecer uma descrição compacta dos dados.

O uso de séries temporais pode ser aplicado para representar diversos eventos de diferentes áreas de conhecimento (AIKES, 2012):

No ramo industrial o uso de ST visa aperfeiçoar a manutenção de equipamentos que sofrem desgaste, com objetivo de maximizar a produção, mantendo-o estável ou até diminuindo os custos de produção;

Na área de astronomia pode ser observada a utilização de ST para o estudo da dinâmica de granulação solar através da obtenção de espectros 2-D;

Na área de estudos ambientais as ST podem ser utilizadas para verificação de estados, funcionamento de ecossistemas e de composição. Outra utilidade da ST é verificar dados de oceanografia, dinâmica da cadeia alimentar, depósitos de carbono, dentre outros;

Na área da saúde ST é largamente utilizada. ST foram analisadas em busca de proporções de diminuição de mortalidade por causas circulatórias e os motivos desta diminuição.

Outras aplicações de modelos de séries temporais incluem a previsão de vendas de vinho tinto, como no exemplo citado anteriormente. Teste de hipóteses, tais como o aquecimento global usando dados de temperatura também são campos largamente estudados usando séries temporais (DAVIS, 2003).

3.1 PREVISÃO DE SÉRIES TEMPORAIS

Segundo Tubino (2006) um Modelo de previsão é dividido em cinco etapas básicas, conforme a Figura 3.

Na primeira etapa, é definido o produto que será previsto, a razão pela qual será realizada a previsão e o grau de confiabilidade; a segunda etapa é realizada a coleta e análise dos dados históricos deste produto, com objetivo de identificar uma técnica de previsão que melhor se adeque a esses dados; no momento em que os dados já foram coletados e analisados, pode-se de fato, definir qual é a técnica de previsão será usada (TUBINO, 2006).

Existem dois tipos de técnicas, sendo elas qualitativas que levam em consideração a opinião de especialistas, e as quantitativas que são baseadas na análise dos dados históricos à partir de modelos estatísticos e matemáticos. Para escolher uma destas técnicas, devem ser levados em consideração os fatores de custo e confiabilidade (TUBINO, 2006).

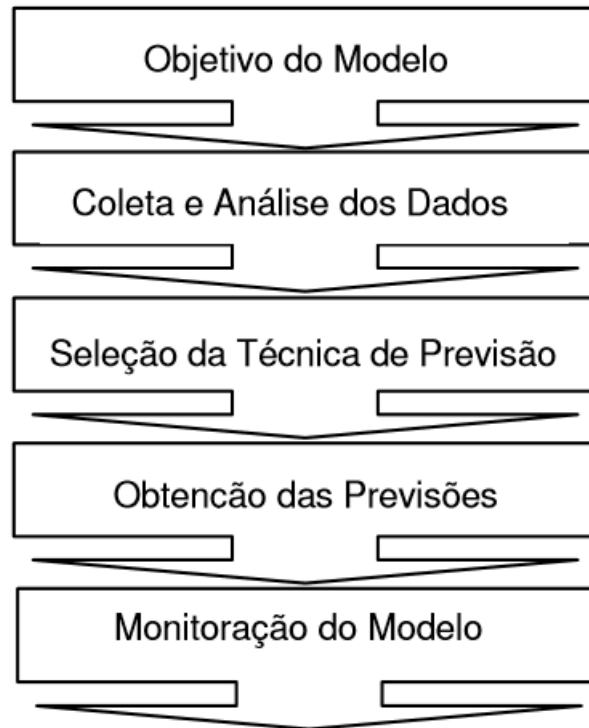


Figura 3 – Modelo de previsão

Fonte: (TUBINO, 2006)

Definido qual técnica de previsão será utilizada, o próximo passo é obter as previsões calculadas pelo modelo e então deve-se compará-las com dados reais no sentido de monitorar o erro obtido, analisando se este erro é satisfatório ou não. Também deve-se monitorar os parâmetros empregados a fim de concluir se os mesmos são válidos ou não (TUBINO, 2006).

Entre as técnicas de previsão de séries temporais utilizadas, têm-se métodos de amortecimento exponencial, médias móveis, regressão linear, projeção de tendência, análise de correlação, dentre outras. O modelo de análise de correlação admite que a previsão está relacionada com algum fator fundamental ou a combinação de fatores no meio ambiente. O Tabela 1 traz uma breve comparação entre algumas das técnicas de previsão.

É sabido que irão ocorrer variações entre os dados reais e a previsão, é pouco provável que o modelo de previsão acerte sempre. Pondo em prática, quase sempre haverá um erro de previsão, e dependendo do tamanho deste erro, os impactos na previsão podem tornar o modelo inútil, sendo necessária alterações nos parâmetros do mesmo até que o erro seja suficientemente satisfatório (FIGUEREDO, 2008).

Muitos esforços são empreendidos na tentativa de aprimorar a precisão da previsão, empregando as melhores técnicas quantitativas e analisando os possíveis cenários, à fim de diminuir os erros. Fundamentalmente a verificação da qualidade das previsões tem propósitos

Técnica	Método Aplicado	Horizonte de Tempo	Complexidade do Modelo	Precisão do Modelo	Dados Necessários
Qualitativa	Delphi	Longo	Alta	Variável	Muitos
Quantitativa Análise de Séries Temporais	Médias Móveis Simples	Curto	Muito baixa	Média	Poucos
	Suavização Exponencial Simples	Curto	Baixa	Adequada	Muito poucos
	Regressão Linear	Longo	Média alta	Média alta	Muitos
Quantitativa Análise de Correlação	Análise de Correlação	Longo	Adequada	Alta	Muitos

Tabela 1 – Quadro de comparação entre técnicas de previsão.

Fonte: (FIGUEREDO, 2008)

econômicos, científicos e administrativos (DAVIS, 2003).

Para Morettin (1981) existem dois importantes componentes em uma série temporal que variam conforme os dados desta série. Os componentes são: sazonalidade e tendência.

- **Sazonalidade** é definida pelos movimentos parecidos nos dados de uma série temporal, tais como número de garrafas de vinhos vendidos, leituras de temperatura. Exemplificando, um empresário que produz e vende vinhos não quer ser pego desprevenido por falta de estoque dos seus produtos, tampouco quer que seus vinhos fiquem estocados por muito tempo, sendo assim, podemos analisar os dados de venda e dizer que vende-se muito mais garrafas de vinho durante uma estação do ano, analisando a sazonalidade dos dados em relação aos meses do ano, pode-se concluir que o vinho é mais vendido no inverno, logo, o empresário poderá prevenir sua produção (GEBHARDKIRCHGÄSSNER, 2013).
- **Tendência** é o movimento nos dados em que se faz de maneira regular e lenta, geralmente representadas por polinômios de baixa ordem. Por exemplo, se o faturamento mensal de um determinado estabelecimento tem aumentado em R\$ 700,00 pode-se então dizer que há uma tendência linear nos dados de faturamento (GEBHARDKIRCHGÄSSNER, 2013).

Merenda (2014) exemplifica graficamente as sazonalidade aditiva e multiplicativa, como também as tendências linear e não linear, como pode-se observar na Figura 4.

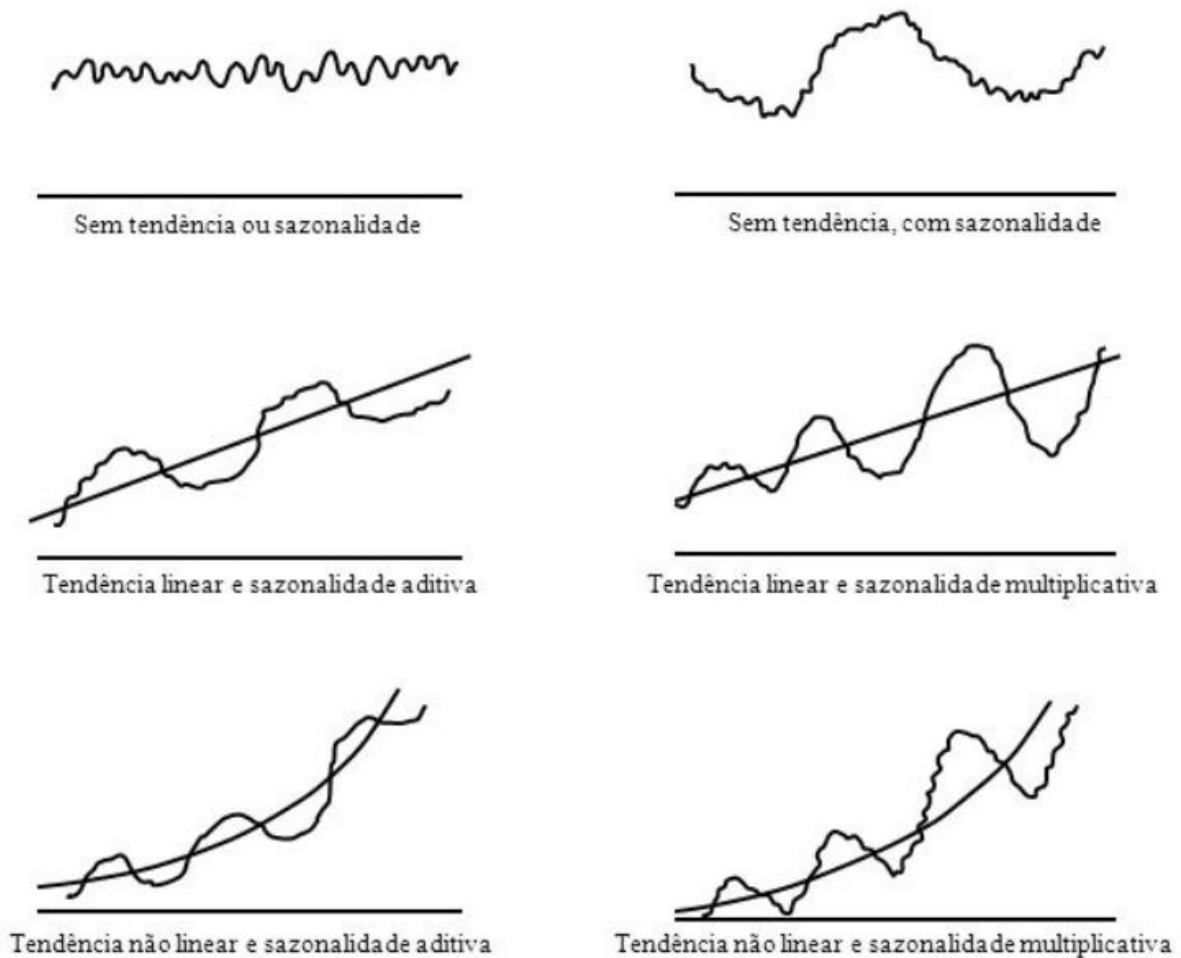


Figura 4 – Sazonalidades e tendências.

Fonte: (MERENDA, 2014)

3.2 DESCOBERTA DE CONHECIMENTO

Instituições que utilizam computadores vêm armazenando dados em grandes volumes, e com uma velocidade crescente. Avanços tecnológicos no armazenamento e mantimento de dados, bem como dispositivos de armazenamento mais rápidos, sistemas de gerenciamento de banco de dados mais eficientes e maior capacidade e menor preço no armazenamento destes dados, têm contribuído significativamente para que existam volumes de dados cada vez maiores. Entretanto, conforme os dados são gerados com maior velocidade e volume, a porção

desses dados que são analisados diminui. Uma parcela pequena de todo esse volume de dados representa a informação potencialmente útil (BATISTA, 2003).

Para Fayyad et al. (1996), o conhecimento é um pequeno grupo de informações que pode adquirido após executar algumas fases do processo, chamado knowledge discovery in databases (KDD). As fases deste processo são: Coleta de dados, pré-processamento, transformação, mineração de dados, avaliação e interpretação. Na Figura 5 pode-se identificar quais são as fases do processo de descobrimento de conhecimento em banco de dados. Em seguida, cada um deles será explanado de forma individual, da maneira que Fayyad et al. (1996) propôs.

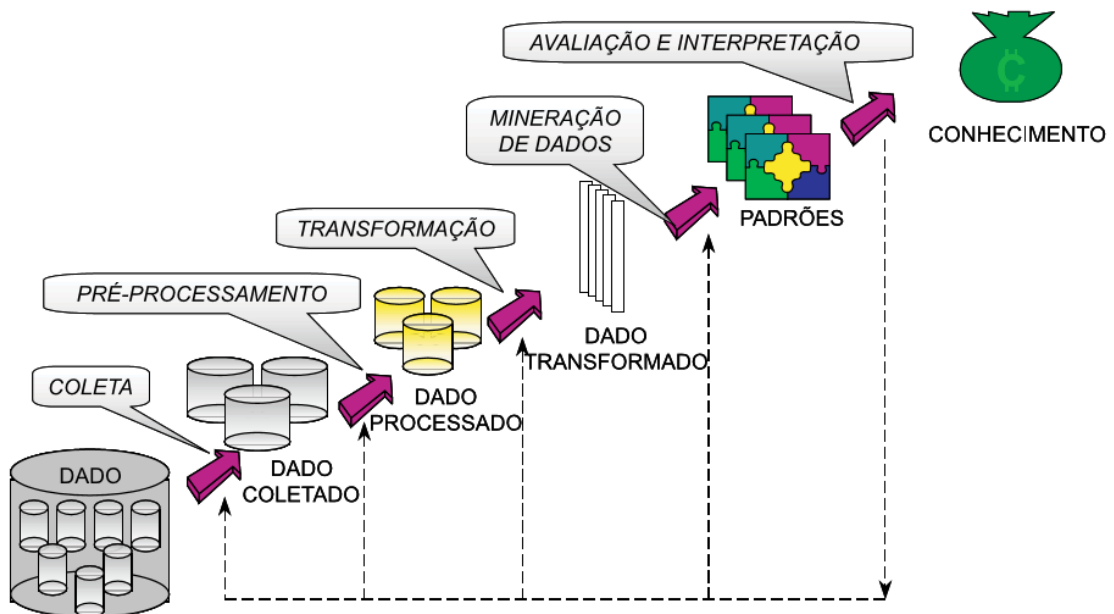


Figura 5 – Processo de descoberta de conhecimento em banco de dados

Fonte: (FAYYAD et al., 1996)

3.2.1 Coleta de dados

Esta fase é muito importante e pode ser uma das mais trabalhosas no processo. Os desafios desta fase podem ser minimizados se a instituição à fornecer os dados possuir um Data Warehouse, que é um repositório de dados construído para dar suporte às pessoas que tomam

decisão em uma determinada empresa, geralmente gerentes e diretores, no sentido de fornecer estatísticas e relatórios das transações realizadas pela empresa. Banco de dados transacionais não são considerados muito adequados para fornecer estatísticas e estratégias para estes gerentes e diretores, sobretudo, frequentemente apresentam problemas de falta de documentação de projeto, integridade de dados e problemas de inconsistência, entre outros (FAYYAD et al., 1996).

Data Warehouse são periodicamente atualizados com os dados dos sistemas transacionais ou dados de fontes externas. Os dados extraídos de diferentes bancos de dados, podem ser integrados e ter sua consistência verificada, antes de serem carregados no Data Warehouse. Portanto, um Data Warehouse pode vir a ser uma boa fonte de dados para um processo de descoberta de conhecimento em banco de dados.

3.2.2 Pré-Processamento

Selecionar os dados é uma das etapas mais importantes do processo, em alguns casos, o pré-processamento é mais importante que a escolha de alguma técnica ou algoritmo que irá processar estes dados. A fase de pré-processamento inicia logo após os dados serem coletados e organizados como um conjunto de dados. Há vários objetivos na fase de pré-processamento. Um dos objetivos é solucionar problemas, tais como identificar dados corrompidos e tratá-los, valores desconhecidos e atributos irrelevantes. Outro objetivo, por exemplo, é alterar a estrutura de dados, através da alteração do grau de granularidade destes dados. São realizadas varias ações no pré-processamento, visando preparar os dados para a próxima fase, que é análise e extração de conhecimento.

De uma maneira geral, o pré-processamento de dados é realizado de maneira semiautomática, pois esta fase depende do conhecimento e da capacidade da pessoa que a conduz em identificar os problemas apresentados nos dados, e utilizar os métodos mais apropriados para solucioná-los. Uma das tarefa a ser realizada no domínio dos dados é a verificação de integridade dos dados. Para verificar estes dados, é frequentemente necessário que quem os faça, tenha conhecimento das restrições de valores aplicáveis a cada atributo. Por exemplo, se uma empresa financeira deseja fornecer crédito a seus clientes, é trivial que este valor de crédito seja obrigatoriamente positivo. Também são comuns valores de integridade entre dois atributos, por exemplo, o valor máximo do seguro de um carro depende do modelo e do ano deste carro. Desta maneira, a partir do conhecimento das restrições de atributos, é

possível utilizar métodos automáticos que encontrem problemas de integridade nestes dados (BATISTA, 2003)

3.2.3 Mineração de dados temporais

O pré-processamento de dados é uma fase do processo de extração de conhecimento que é necessário grande conhecimento de domínio. Para muitos pesquisadores, os dados que são coletados diretamente em bases de dados podem ser de má qualidade, pois possuem informações imprecisas e incorretas, podem ter uma grande quantidade de valores desconhecidos, como também diferentes tipos de dados.

A grande diferença entre os dados temporais e os não temporais está na ordenação, onde a amostra de dados não temporais não se preocupa com a ordem e ocorrem de forma aleatória, já nos dados temporais é preservada essa ordem, sendo assim, a forma que se preparam estes dados é diferente. Segundo Pyle (1999), na fase de pré-processamento de dados, temos algumas informações que devem ser levadas em consideração para preparação dos dados, sendo elas:

- **Outliers**, são os dados que fogem do padrão, podendo ou não estes dados estarem corretos. Quando corretos, os *outliers* podem ocorrer devido algum evento externo incomum, por exemplo, o registro de 38°C em Curitiba no mês de Junho. Já quando não estão corretos, os *outliers* ocorrem devido à um erro de entrada de dados, por exemplo, uma balança falhou e aferiu 200Kg para um recém nascido. *Outliers* podem ser resolvidos simplesmente removendo-os ou mudando os seus valores para os mesmos dos índices anteriores, caso essas mudanças não afetem o modelo.
- **Valores faltantes**, podem ocorrer de duas formas, sendo elas valores faltantes de índice temporal que normalmente são tratados utilizando o próximo valor do índice, e valores faltantes do dado, tendo maior complexidade para resolver este segundo. Uma possível maneira de resolver estes valores faltantes é completando estes valores baseado em algum padrão existente na série, por exemplo, repetir o valor de índice anterior ou posterior.
- **Amostragem irregular**, na maioria das séries temporais, a sua mensuração é realizada através de intervalos idênticos de amostragem, porém quando isto não ocorre é necessário realizar uma transformação na série, afim de ajustar os dados e seus valores.

De maneira semelhante à mineração de dados, o objetivo da mineração de dados temporais é identificar relações ocultas entre as sequências de eventos (AIKES, 2012).

3.2.4 Transformação de Dados

Nesta fase, o principal objetivo é transformar a representação dos dados para que seja superado qualquer tipo de limitação existente nos algoritmos que serão utilizados para a extração de conhecimento. A decisão de quais transformações serão necessárias vai depender de qual algoritmo será utilizado na fase de mineração. Segundo Batista (2003) algumas das transformações mais comuns são:

- **Normalização:** A normalização consiste em modificar os intervalos dos atributos, por exemplo, podemos transformar o intervalo [1,15] para o [1,3], de maneira que cada 5 valores do primeiro intervalo, seja transformado em apenas um valor. Este tipo de transformação é valiosa especialmente para os algoritmos que calculam distâncias entre atributos. Dentro das distâncias utilizadas tem-se: Distância euclidiana, distância euclidiana média, distância de Mahalanobis, distância de Manhattan, dentre outras;
- **Discretização de atributos quantitativos:** Muitos algoritmos utilizados na fase de mineração têm a limitação de só trabalhar com atributos qualitativos, porém muitos conjuntos de dados só possuem atributos quantitativos, sendo assim, é necessário aplicar algum método que transforma um atributo quantitativo em qualitativo, ou seja, transformar em faixas de valores;
- **Transformação de atributos qualitativos em quantitativos:** Existem algoritmos que não são capazes de processar atributos qualitativos, sendo assim, é necessário a conversão de atributos qualitativos para quantitativos. De maneira geral, atributos qualitativos tal como azul, verde, amarelo, podem ser transformados em números, como 1, 2 e 3. Porém esta transformação pode criar uma ordem nos valores que não é real, pois analogamente, tem-se que azul está mais perto do verde, do que do amarelo, entretanto esta analogia não faz sentido algum no mundo real. Atributos qualitativos com ordem, bem como leve, médio e pesado, podem ser facilmente mapeados para valores numéricos, sem perder a ordem;
- **Atributos de tipos de dado complexos:** Grande parte dos algoritmos utilizados para extrair padrões não possui a habilidade de trabalhar com tipos de dados complexos, por exemplo, atributos do tipo data e hora não são comumente analisados pela maioria dos algoritmos. Sendo assim, precisa-se converter estes atributos de tipo complexo para outro tipo de dado, para que o algoritmo possa reconhecê-los. No específico caso de data e hora, pode-se utilizar a conversão para o tipo inteiro, calculando-se a diferença em dias, meses, ano ou qualquer unidade de tempo, entre as datas do atributo em questão e uma data fixa

qualquer. Por exemplo, pode-se calcular a idade de uma pessoa, em anos, calculando a diferença do ano em que esta pessoa nasceu e o ano da data atual.

De maneira geral, as transformações são realizadas com objetivo de reduzir a dimensão da série, buscando perder a menor quantidade de informação relevante possível; e isolar características específicas dos seus componentes (AIKES, 2012).

3.2.5 Avaliação e interpretação

Após a fase de mineração dos dados é necessário avaliar e interpretar os resultados obtidos. Esta fase engloba todos os participantes. O analista de dados procura descobrir se o classificador atingiu as expectativas, através da avaliação dos resultados de acordo com métricas, sendo as mais comuns taxa de erro, complexidade do modelo e tempo de processamento. O especialista no domínio verifica a compatibilidade e os resultados obtidos com o conhecimento disponível do domínio. No final, é o usuário que julga a aplicabilidade dos resultados.

No processo de obtenção de conhecimento, é esperado que o resultado final seja compreensível, porém a definição de compreensibilidade não é muito trivial. Em alguns contextos, pode ser fácil aplicar uma métrica para compreensibilidade, como por exemplo uma árvore de decisão, que pode ser calculado o grau de compreensibilidade através do número de nós que esta árvore possui. Entretanto, até o momento não existe uma maneira efetiva de medir compreensibilidade do conhecimento.

4 INTELIGENCIA ARTIFICIAL

O nome da espécie é HOMO SAPIENS (homem sábio) devido a inteligência ser tão importante. Por milhares de anos, busca-se entender o pensar. O ramo da inteligência artificial, ou Artificial Intelligence (AI), vai ainda mais além, ele busca não apenas compreender, como também construir entidades inteligentes (RUSSELL, 2010).

Até alguns anos atrás, a AI era muitas vezes vista como uma área mais teórica, aplicada apenas em pequenos problemas curiosos, mas de pouco valor prático. Muitos problemas práticos da vida real que necessitavam da computação para se ter solução, eram resolvidos à partir de codificação em alguma linguagem de programação. Com o passar dos anos houve maior disseminação do uso de algumas técnicas de AI para solução de problemas reais. Frequentemente esses problemas eram trazidos computacionalmente por meio da aquisição de conhecimento de especialistas de um domínio, por exemplo, um domínio da meteorologia, que era então codificado por regras lógicas em um programa de computador. Estes programas eram conhecidos como Sistemas Baseados em Conhecimento ou como Sistemas Especialistas (FACELI et al., 2011).

O processo para adquirir o conhecimento normalmente envolvia entrevistas com os especialistas, a fim de descobrir quais regras eles utilizavam para a tomada de decisão. Esse processo de obtenção de conhecimento possuía várias limitações, como subjetividade, decorrente do uso da intuição do especialista na tomada de decisão, bem como, falta de cooperação por parte do especialista, por causa do medo de ser dispensado após repassar o conhecimento a ele requerido (FACELI et al., 2011).

Nos últimos anos, com o crescente volume de dados gerados por diferentes setores e a crescente complexidade dos problemas a serem tratados computacionalmente, tornou-se fundamental o uso de ferramentas computacionais mais sofisticadas, capazes de reduzir a necessidade de intervenção humana e da dependência de especialistas. Para isso, essas técnicas e ferramentas deveriam ser capazes de criar, a partir da experiência passada, uma função ou hipótese que seja capaz de resolver um determinado problema. Por exemplo, um determinado produto é lançado no mercado e os gerentes de venda desejam saber qual é o público alvo deste produto, baseado nos dados de compras dos clientes dos supermercados. Este processo de

indução de uma hipótese a partir da experiência passada é chamado de Aprendizado de Máquina (FACELI et al., 2011).

Hoje em dia, a AI engloba uma enorme variedade de subcampos. De uma maneira geral é muito versátil e muito utilizada nas áreas de aprendizagem e percepção, mas pode também abranger até tarefas mais específicas, como demonstração de teoremas matemáticos, jogos de xadrez, direção de um carro em uma estrada pavimentada, diagnósticos de doenças, entre outros (RUSSELL, 2010).

Considerada essencial para um comportamento inteligente, a capacidade de aprendizado é capaz de realizar atividades como memorizar, melhorar habilidades motoras e cognitivas, observar e explorar situações para que aprenda os fatos. No aprendizado de máquina, computadores podem ser programados para aprender a partir da experiência passada. Para isso, empregam o princípio da inferência, chamado de indução, que gera conclusões genéricas com base em um conjunto particular de dados. Desta forma, algoritmos de aprendizado de máquina aprendem a induzir uma hipótese que é capaz de resolver um problema baseado em dados que representam as instâncias desse problema a ser resolvido (FACELI et al., 2011).

Apesar do fato de Aprendizado de Máquina muitas vezes ser associado à Inteligência Artificial, existem outras áreas que também são importantes e têm várias contribuições significativas no avanço do Aprendizado de Máquina, tais quais, Teoria da Computação, Teoria da Informação, Neurociência, Probabilidade e Estatística, dentre outras. São muitos os algoritmos e suas variações para utilização de Aprendizado de Máquina, além disso, todo instante surgem novas variações de problemas reais já tratados. Ainda segundo (FACELI et al., 2011), várias aplicações foram bem-sucedidas com técnicas de Aprendizado de Máquina para solução de problemas reais, dentre elas, podem ser citadas:

- condução autônoma de automóveis em rodovias pavimentadas;
- predição de taxas de cura de pacientes com diferentes doenças;
- reconhecimento de palavras na língua falada;
- detecção de fraude em cartões de crédito;
- diagnósticos de câncer por meio de análise de dados de expressão genética;
- ferramentas que jogam xadrez e gamão de forma parecida a campeões.

4.1 ABORDAGENS NO ESTUDO DE INTELIGÊNCIA ARTIFICIAL

Historicamente, tem-se quatro estratégias a serem seguidas para o estudo de AI, sendo elas: pensar como humano, pensar racionalmente, agir como humano, agir racionalmente. Uma abordagem focada nos seres humanos deve ser em parte um conhecimento empírico, que envolvem hipóteses e confirmação experimental. Uma abordagem racionalista deve envolver uma combinação de engenharia e matemática. Cada uma delas tem ao mesmo tempo desacreditado e ajudado a outra. Abaixo será explanado cada uma individualmente segundo (RUSSELL, 2010).

4.1.1 Agindo de forma humana

O teste de Turing, proposto por Alan Turing (1950), foi desenvolvido para fornecer uma definição de inteligência. Se um interrogador humano, depois de propor algumas perguntas por escrito, não conseguir distinguir se as respostas vêm de um ser humano ou de um computador, então o computador que respondeu estas perguntas passou no teste. Pode-se observar que programar o computador para passar no teste de Turing exige muito trabalho, pois o computador necessitaria ter as seguintes capacidades (RUSSELL, 2010):

- processamento de linguagem natural, para possibilitar que o computador se comunique com sucesso com o interrogador;
- representação de conhecimento, para armazenar o que sabe ou ouve;
- raciocínio automatizado, para usar as informações armazenadas com o objetivo de responder as perguntas e tirar conclusões;
- aprendizado de máquina, para se adaptar a novas circunstâncias e para detectar e extrapolar padrões.

4.1.2 Pensando de forma humana

Se o que pretende-se é dizer que dado programa pensa como um ser humano, precisa-se determinar como os seres humanos pensam, para se poder fazer esta comparação. Existem três modos de fazer isso: através de experimentos psicológicos – observando uma pessoa em ação; através da introspecção – buscando captar os próprios pensamentos à medida que são criados; e através de imagens cerebrais – observando o cérebro em ação. Após se obter uma teoria da mente abastadamente precisa, será possível expressar essa teoria como um programa computacional. Caso os comportamentos analisados e a sincronização do programa forem semelhantes, é possível dizer que alguns dos mecanismos do programa podem também estar operando nos seres humanos (FACELI et al., 2011; RUSSELL, 2010).

4.1.3 Pensando racionalmente

Aristóteles foi um dos primeiros a tentar entender o “pensamento correto”, ou seja, os processos de raciocínio irrefutáveis. Seus silogismos forneceram padrões para estruturas de argumentos, que incessantemente terminavam em conclusões corretas ao perceberem as premissas corretas. Por exemplo, “Jorge é um homem; todos os homens são mortais; logo, João é mortal”. Essas leis deveriam governar a operação da mente. Seus estudos deram início ao importante campo chamado lógica (RUSSELL, 2010).

Por volta de 1965, haviam programas computacionais que podiam, originalmente, resolver qualquer problema solucionável descrito em notação lógica. Contudo, se não houvesse solução este programa poderia entrar em loop infinito. Esta abordagem de lógica enfrenta dois principais obstáculos. Primeiro deles é que não é fácil enunciar o conhecimento informal nos termos formais, principalmente quando o conhecimento é menos de 100% correto. Por segundo lugar, existe uma grande diferença entre ser capaz de resolver problemas e, de fato, resolve-los. Mesmo problemas com poucos fatos podem esgotar os recursos computacionais (FACELI et al., 2011; RUSSELL, 2010).

4.1.4 Agindo racionalmente

A palavra agente vem do latino *agere*, que significa fazer. Um agente é simplesmente algo que age. Provavelmente todos os programas computacionais realizam alguma tarefa, mas espera-se que um agente computacional faça mais que uma tarefa, espera-se que ele opere sob controle autônomo (automaticamente), perceba seu ambiente, adapte-se a mudanças, persista por um período de tempo prolongado e seja capaz de criar e perseguir metas (FACELI et al., 2011; RUSSELL, 2010).

A abordagem do agente racional tem duas vantagens sobre quaisquer outras abordagens. Primeiramente ela é mais geral que a abordagem das “leis do pensamento”, pois a inferência correta é apenas um dos vários mecanismos possíveis para alcançar a racionalidade. Em segundo lugar, ela é mais acessível cientificamente do que as estratégias baseadas no comportamento ou no pensamento humano. Matematicamente falando, o padrão de racionalidade é bem definido e completamente geral, podendo ser desempacotado para gerar os modelos de agente que comprovadamente poderão atingi-lo (FACELI et al., 2011),(RUSSELL, 2010).

4.2 REDES NEURAIS

O primeiro trabalho sobre Redes Neurais foi desenvolvido no ano de 1943. Neste, foi desenvolvido um estudo sobre o comportamento do neurônio biológico, com intuito de criar um modelo matemático para o mesmo. Foram de extrema importância as conclusões desta pesquisa para futuras implementações computacionais. Segundo estas conclusões, a atividade do neurônio é do tipo tudo ou nada, e a atividade de qualquer sinapse inibitória, pode prevenir a excitação do neurônio naquele momento (CARDON, 1994).

O que a primeira afirmação quer dizer é que um neurônio estará no estado de ativo se a sua saída ultrapassar o valor limite, senão, ficará no estado de repouso. Este principio chama-se função limiar, e atingir esta função limiar significa transmitir a saída a outros neurônios da rede. Analogamente, podemos pensar que a função limiar seja um número a ser atingido, por exemplo, a soma das entradas precisa ser maior que 20. Caso a soma das entradas não atinja o valor 20, este neurônio não irá transmitir sinal algum para os neurônios da próxima camada. A segunda afirmação quer dizer que todo neurônio terá entradas de valor positivo e negativo,

esta entrada deverá inibir a saída. Se esta entrada tem um valor positivo, a saída será excitada (CARDON, 1994).

O cérebro humano contém cerca de 100 bilhões neurônios. Cada neurônio processa e se comunica com milhões de outros continuamente e paralelamente. A estrutura é constituída a partir de neurônios biológicos. Estes são divididos em três seções: o corpo da célula, axônio e os dendritos, cada um com suas funções específicas, porém são funções complementares. Os dendritos têm a função receber as informações, os chamados impulsos nervosos, provenientes de outros neurônios e levados até o corpo celular. No corpo celular, a informação é processada e novos impulsos são gerados. Estes novos impulsos criados são transmitidos, passando do axônio até os dendritos dos neurônios da próxima camada. Essa transmissão de uma terminação axônica de um neurônio com o dendrito de outro neurônio, é chamada de sinapse. Através destas sinapses que os neurônios são unidos e formam uma rede neural. As sinapses funcionam como transmissoras de impulsos entre os neurônios de uma rede. Na Figura 6 pode-se ver estes componentes de um neurônio biológico (BRAGA ANDRE PONCE DE LEON F. DE CARVALHO, 2007).

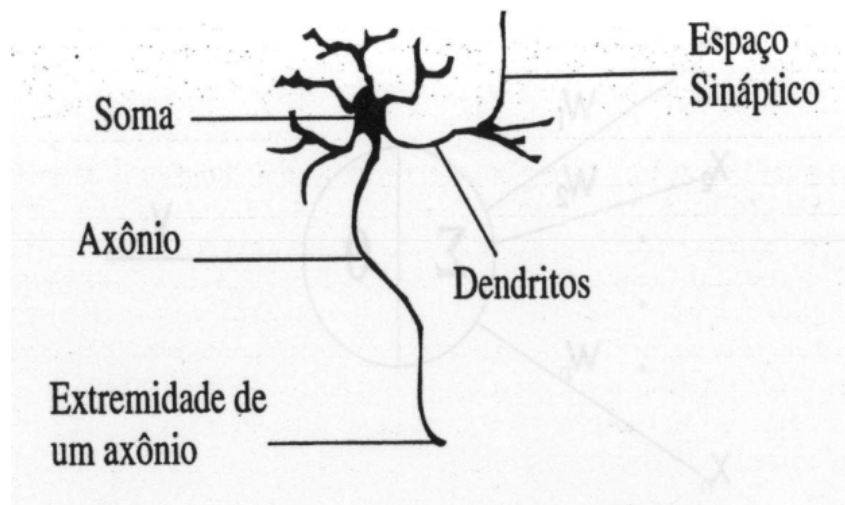


Figura 6 – Neurônio Biológico.

Fonte: (BRAGA ANDRE PONCE DE LEON F. DE CARVALHO, 2007)

As RNA tem uma estrutura paralelamente distribuída e a habilidade de aprender e generalizar, por este fato, ela pode ser utilizada para resolução de problemas complexos e de grande escala que geralmente são intratáveis. Uma RNA precisa trabalhar em conjunto, pois problemas grandes e complexos são compostos por tarefas menores e relativamente fáceis de serem resolvidos. Segundo Haykin (2000) o uso das RNA oferecem as seguintes propriedades úteis e capacidades:

- **Não-linearidade:** O conjunto de neurônios artificiais que formam uma rede tem a

habilidade de resolver problemas não lineares. Esta propriedade é muito importante, particularmente quando os sinais de entrada também não forem lineares, como por exemplo o reconhecimento de voz;

- **Mapeamento de Entrada-Saída:** A partir do paradigma de aprendizagem supervisionada, os pesos sinápticos de uma RNA são atualizados afim de minimizar a diferença entre a resposta desejada e a resposta da rede. Desta forma, no treinamento da RNA são dadas uma entrada específica e espera-se uma saída, se esta saída não for a desejada, muda-se os pesos sinápticos;
- **Adaptabilidade:** Uma RNA tem a capacidade de adaptar seus pesos sinápticos a modificações externas. Uma RNA treinada para atuar em um ambiente específico pode ser facilmente retreinada para algumas modificações nas variáveis deste mesmo ambiente;
- **Resposta a Evidências:** Quando utilizada como classificador, uma RNA pode fornecer informação sobre qual padrão foi escolhido, bem como qual a confiança ou crença nessa decisão tomada;
- **Informação Contextual:** Uma RNA é um algoritmo preditivo, o conhecimento é representado pela sua própria estrutura e estado de ativação. Cada neurônio artificial é potencialmente afetado pela atividade de todos os outros neurônios da mesma rede;
- **Tolerância a Falhas:** Pelo fato da RNA ter uma natureza distribuída da informação armazenada, pequenos danos aos neurônios que fazem parte da rede, podem ser facilmente tolerados, uma vez que a estrutura global da rede não é afetada de forma massiva;
- **Implementação em *Very Large Scale Integration* (VLSI):** A natureza maciçamente paralela de uma RNA faz com que ela seja rápida na computação de determinadas tarefas, se tornando ideal para implementação utilizando a tecnologia de integração em escala muito ampla ou VLSI;
- **Uniformidade de Análise e Projeto:** Os neurônios artificiais basicamente representam um ingrediente comum a todas as RNAs, esta uniformidade torna possível compartilhar teorias e algoritmos de aprendizagem em diferentes projetos ou áreas de estudo;
- **Analogia Neurobiológica:** Uma RNA pode ser utilizada como ferramenta de pesquisa para interpretar fenômenos neurobiológicos, pelo fato de uma RNA ser um projeto análogo ao cérebro.

4.2.1 Perceptron

O Perceptron se trata basicamente de uma RNA de única camada. Geralmente uma rede Perceptron é constituída por apenas um neurônio artificial e um polarizador. Pelo fato de ela ter apenas uma camada oculta, este tipo de RNA só conseguem classificar padrões que sejam linearmente separáveis, ou seja, classes linearmente separáveis em um hiperplano. Segundo Vellasco (2007) as características básicas de um Perceptron são:

- função de ativação: degrau;
- topologia: uma única camada de neurônios;
- algoritmo de aprendizado: supervisionado.;
- valores entrada/saída: binários [-1, +1].

Como no mundo real, a maioria dos padrões não são linearmente separáveis, o uso do Perceptron não se mostra muito eficaz. Porém para problemas não linearmente separáveis, pode-se usar uma variação do Perceptron, chamada Multilayer Perceptron (Perceptron de múltiplas camadas), que será apresentado a seguir.

4.2.2 Multilayer Perceptron

As RNAs multi camadas tipicamente são constituídas por uma camada de entrada, uma ou mais camadas escondidas ou camadas ocultas, e uma camada de saída. Os sinais transmitidos pelos neurônios são propagados para frente, camada por camada. O treinamento das RNA multi camadas é feito de forma supervisionada e utiliza-se o algoritmo Backpropagation (retropropagação de erro) que será apresentado posteriormente. Segundo Vellasco (2007), existem três características básicas das RNAs multi camadas:

- O modelo utiliza uma função de ativação não-linear. A não linearidade é do tipo suave, ou seja, é diferenciável em qualquer ponto. Um modelo que fosse somente linear, teria a relação entrada-saída reduzida, existente em RNAs de única camada;
- O modelo possui uma ou mais camadas ocultas, além das suas camadas de entrada e saída. Os neurônios da camada oculta possibilitam que a rede aprenda tarefas complexas, extraindo progressivamente as características mais significativas das entradas;
- Uma rede multi camadas possui alto grau de conectividade determinado pelos seus pesos,

qualquer modificação nessa conectividade requer alteração dos pesos.

Nas RNAs Multilayer Perceptron o algoritmo de treinamento mais difundido é o Backpropagation. Este algoritmo consiste basicamente em calcular o erro na saída da rede e retropropagá-lo pela rede. A maneira que o algoritmo atua é atualizando os pesos sinápticos, camada a camada com o objetivo de diminuir o erro da próxima saída (VELLASCO, 2007).

No Backpropagation os pesos sinápticos são atualizados baseando-se no método de gradiente descendente. A ideia central deste método é realizar modificações proporcionais ao gradiente do erro, o método utiliza a primeira derivada de todas as inclinações. Em outras palavras, os pesos que sofrerão maior impacto no momento da atualização, serão os pesos que tiveram maior influência no erro da saída (VELLASCO, 2007).

4.3 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentadas definições de inteligência artificial, definições de agir e pensar de forma humana e racionalmente. Foram apresentadas aplicações de redes neurais e como elas funcionam. O próximo capítulo apresenta os materiais e métodos utilizados no desenvolvimento deste trabalho.

5 MATERIAIS E MÉTODOS

O grande volume de dados sobre os valores das ações não significam nada se não forem tratados e analisados. Neste capítulo serão expostas ferramentas e métodos que irão auxiliar no tratamento e análise dos dados e irão retornar um resultado mais compreensível. Para o tipo de pesquisa teórica, foram utilizadas como principais palavras-chave: Redes neurais artificiais, previsão, ativo financeiro.

5.1 COLETA DE DADOS

Os dados financeiros necessários para a realização deste trabalho foram basicamente a cotação de uma ação da empresa Tesla Motors Inc. – TSLA, extraídos da bolsa de valores NASDAQ. Para que houvesse uma quantidade considerável de registros, os dados foram coletados a cada 5 minutos durante o período de 10 dias. Os dados só puderam ser coletados durante o intervalo de abertura e fechamento da bolsa NASDAQ, que é dado das 10h30min as 17h00min (horário de Brasília). O critério de coleta de dados inicial era que a frequência de coleta fosse de minuto a minuto, porém após uma análise destes dados, notou-se que a variação do valor dos ativos era muito pequena e que se alterada para cada 5 minutos, haveria uma variação mais expressiva.

A ferramenta utilizada para coletar os dados foi uma application programming interface(API) da Google chamada Google Finances¹. Nesta API é possível coletar dados históricos e dados em tempo real de diversas empresas nacionais e internacionais a partir de um endereço inserido no navegador, como visto na Figura 7. Os parâmetros da API são: Frequência (em segundos), quantidade de dias e a sigla da empresa na bolsa de valores. Para este trabalho, como dito no bloco anterior, foram coletado dados históricos de 10 dias a cada 5 minutos da TSLA, no período de 01/02/2017 a 14/02/2017 (de segunda a sexta feira).

¹<https://www.google.com.br/finance/>



```

EXCHANGE%3DNASDAQ
MARKET_OPEN_MINUTE=570
MARKET_CLOSE_MINUTE=960
INTERVAL=300
COLUMNS=DATE,CLOSE,HIGH,LOW,OPEN,VOLUME
DATA=
TIMEZONE_OFFSET=-240
a1493731800,324.02,324.1,324,324,73303
1,325.23,327.6599,324.02,324.1,296594
2,322.8601,326.27,322.85,325.2,238025
3,322.133,323.8599,321.33,323.22,212527
4,320.53,322.1735,320.5,322.015,176468
5,321.2878,322.45,320.3,320.51,143744
6,320.89,321.6,320.61,321.17,100498
7,321.15,321.6,320.67,320.72,114104
8,320.6991,321.42,320.01,321.155,122136
9,320.1,321.3,319.34,320.7749,154584
10,320.56,320.84,319.81,319.925,117021
11,320.62,321.19,320.5945,320.625,84984

```

Figura 7 – Google Finance

Fonte: Autoria própria

Para este trabalho, foram criados dois grupos primários para realização dos treinamentos e testes. Em um dos grupos, somente a cotação do momento inicial dos 5 minutos foram analisados, no segundo grupo, cada registro financeiro continha os valores de:

- Cotação abertura, representado pela letra O;
- Mínima (cotação mínima deste período), representado pela letra M;
- Máxima (cotação máxima deste período), representado pela letra H;
- Cotação fechamento, representado pela letra C;
- Volume negociado, representado pela letra V.

5.2 FORMATO DOS DADOS

Os dados obtidos pela API do Google Finances é dado por um arquivo do tipo *comma separated value* (CSV) corrido, na forma de uma série temporal, onde cada linha representa uma instância de 5 minutos ao decorrer dos 10 dias que foram requisitados. Estes dados foram colados em um documento no formato Folha de Cálculo do Microsoft Excel (.xlsx). Em seguida na guia Dados do Excel, foi utilizado o recurso Texto para Colunas, separando os dados em

colunas. Após algumas transformações e tratamento dos dados, que serão vistas a seguir, os dados são salvos no formato xlsx, e então é realizada a conversão para CSV novamente. Para esta conversão, foi utilizado um conversor online Convertio², pois houveram alguns problemas no processamento da RNA quando os dados eram convertidos diretamente em CSV a partir do Excel.

5.3 AMBIENTE

Para a realização deste trabalho, foi inicialmente utilizado um microcomputador com processador AMD FX 8120 octa-core 3.1ghz, 4GB de memória RAM. Porém no decorrer deste trabalho, se viu necessário um upgrade de memória RAM, com o objetivo de reduzir o tempo de treinamento e testes do modelo. Alguns treinamentos/testes foram realizados com 4GB de RAM e outros com 8GB de RAM. Foi certificado que os dados de saída do processamento da RNA não seriam alterados pelo aumento de memória RAM no sistema. Único fator alterado nos testes realizados foi o tempo de processamento.

Neste microcomputador foi utilizado o sistema operacional Windows 10 64bits³, também foram utilizados os softwares Google Chrome⁴, Excel 2010⁵ e o Weka⁶.

5.4 WEKA

Waikato Environment for Knowledge Analysis (WEKA) é uma suíte de aprendizado de máquina, escrita em Java, desenvolvida pela University of Waikato, Nova Zelandia. É um software livre licenciado sob a General Public License (GPL). A Weca é uma suíte que contém uma coleção de ferramentas de visualização e algoritmos para análise de dados e modelagem preditiva, juntamente com interfaces gráficas para facilitar o acesso a estas funções. A Weka

²<https://convertio.co/pt/>

³<https://www.microsoft.com/pt-br/software-download/windows10>

⁴<https://www.google.com.br/chrome/browser/desktop/>

⁵<https://www.microsoft.com/pt-br/download/details.aspx?id=28534>

⁶<https://sourceforge.net/projects/weka/>

suporta várias tarefas para mineração de dados, mais especificamente pré-processamento de dados, agrupamento, classificação, regressão, visualização e seleção de atributos.

5.4.1 Interface do usuário

Interface do usuário no Weka se dá através do Explorer, mas as funcionalidades também podem ser acessadas através de linhas de comando. Há também o Experimenter, que permite comparar o desempenho de algoritmos preditivos. Há 6 painéis no Explorer para navegação, sendo eles:

- O painel Preprocess** tem facilidades para importar dados de um base de dados, arquivos no formato CSV, entre outros. Também é utilizado para pré-processamento de dados usando um algoritmo de filtragem. Estes filtros podem ser usados para transformar os dados (por exemplo, transformando atributos numéricos em discretos) e permite excluir atributos e instâncias de acordo com critérios específicos;
- O painel Classify** permite a aplicação de algoritmos de classificação e regressão para o conjunto de dados resultante, a fim de estimar a precisão do resultado do modelo preditivo e visualizar predições erradas;
- O painel Associate** permite o acesso a regras de aprendizado que tentam identificar todas relações importantes entre os atributos nos dados;
- O painel Cluster** dá acesso às técnicas de agrupamento no Weka, o algoritmo K-MÉDIAS é um exemplo de técnica de agrupamento suportada pelo Weka;
- O painel Select Attributes** fornece algoritmos para identificar os atributos mais preditivos em um conjunto de dados;
- O painel Visualize** mostram gráfico de dispersão da matriz, gráficos individuais podem ser selecionados, ampliados e analisados utilizando vários operadores de seleção. Neste trabalho será utilizada a versão estável mais recente 3.8.0 do Weka. A escolha desta ferramenta foi motivada pelo fato de já ter tido contato com a mesma durante a graduação, se mostrando muito eficiente e didática nas atividades desenvolvidas.

5.4.2 MultiLayer Perceptron

É uma modificação do padrão Perceptron e pode distinguir os dados que não são linearmente separáveis. Como visto anteriormente, o Weka possui ferramentas para pré-processamento de dados, classificação, regressão, clustering, regras de associação e visualização. A RNA está categorizada como um algoritmo de classificação. Neste trabalho foi utilizado a RNA Multilayer Perceptron que pode ser selecionado dentro da aba Classify no software Weka.

5.4.3 Parametros de Rede Neural Artificial no Weka

Os parâmetros para o algoritmo MultilayerPerceptron considerados pelo Weka são:

- Graphical user interface (GUI) ou interface gráfica do usuário assume apenas dois valores, verdadeiro ou falso. Quando verdadeiro traz uma interface da rede neural durante o treinamento, sendo possível fazer uma pausa e alterar a RNA;
- autoBuild, assume apenas dois valores, verdadeiro ou falso, quando verdadeiro faz com que todas as conexões entre as camadas ocultas sejam criadas de forma automática;
- debug, assume apenas dois valores, verdadeiro ou falso, quando verdadeiro a saída gerada traz informações a mais;
- decay, assume apenas dois valores, verdadeiro ou falso, quando verdadeiro faz com que a taxa de aprendizagem da RNA diminua, isso irá dividir a taxa de aprendizagem podendo ajudar a rede a não divergir a sua saída bem como a performance geral;
- hiddenLayers, esse parâmetro só terá efeito se o autoBuild estiver com o valor verdadeiro. É definida a quantidade de camadas ocultas da RNA, começando com 0, caso definido nenhuma camada escondida. Existem ainda os valores pré-definidos sendo eles:
 - $a = (\text{attribution} + \text{classes}) / 2$, faz uma média de atributos e classes;
 - $i = \text{attribution}$, a quantidade de atributos é a quantidade de camadas escondidas;
 - $o = \text{classes}$, a quantidade de classes é a quantidade de camadas escondidas;
 - $t = \text{attribution} + \text{classes}$, a soma dos atributos e classes é a quantidade de camadas;
- learningRate, valor que os pesos sinápticos serão atualizados;
- momentum, impulso aplicado aos pesos sinápticos durante a atualização;
- nominalToBinaryFilter, Converte todos os atributos nominais em atributos numéricos

binários;

- `normalizeAttributes`, Normaliza todos os valores numéricos no conjunto de dados determinado;
- `normalizeNumericClass`, Normaliza a classe se for numérico. Isso ajuda a melhorar o desempenho da rede, Normaliza a classe para estar entre -1 e 1;
- `reset`, assume apenas dois valores, verdadeiro ou falso, se verdadeiro permite que a RNA reinicie com uma taxa de aprendizagem menor quando divergir da resposta e o treinamento ocorre de novo;
- `seed`, utiliza números randômicos para definir os pesos sinápticos iniciais entre as conexões e também para embaralhar os dados de treinamento;
- `trainingTime`, o número de ciclos que a RNA irá executar;
- `validationSetSize`, porcentagem de tamanho do conjunto de validação, se definido zero, nenhum conjunto de validação será utilizado;
- `validationThreshold`, utilizado para contar quantas vezes seguidas ocorrem erros, para que o treino possa ser encerrado.

Antes de executar o treinamento da RNA é necessário configurar os parâmetros, a melhor configuração possível somente é conhecida depois de efetuado testes com diferentes valores nos parâmetros, sendo que ao final de cada teste são exibidos os resultados. O valor utilizado como critério para avaliar se a configuração foi boa ou não, foi o Mean absolute error, que define a média das somas do erro de cada instância de teste que a RNA executou.

Para este trabalho 126 diferentes configurações foram executadas. Através da busca em grade, os parâmetros são modificados e testados, para que se escolha quais deles obtiveram melhor resultado. Os parâmetros que foram alterados a fim de escolher a configuração foram:

- `learningRate`, fixo em 0.01;
- temporalidade, variando o intervalo de 15 até 120 minutos;
- `trainingTime`, variando de 500 até 5000 vezes;

Para este trabalho, foram testados algumas variações destes parâmetros, com diferentes grupos de dados financeiros da Tesla, com o objetivo de minimizar o erro do modelo. Após alguns testes, notou-se que para diferentes grupos de testes, o valor 0.01 para o parâmetro `learningRate` tinha o melhor resultado para todos os grupos de dados testados. Bem como, o melhor valor para o parâmetro `momentum` e `hiddenLayers` foi o padrão do Weka, que é 0.2 e $\frac{a}{2}$ (atributos+classes)/2 respectivamente.

Como o algoritmo `MultilayerPerceptron` no Weka não trata temporalidade, neste trabalho foram utilizados alguns recursos para que a série temporal fosse representada por colunas de cada instância, divididas em intervalos de 5 minutos. Em uma sessão posterior,

será melhor explicado como foram divididas as séries temporais para realização dos testes.

O parâmetro que mais variou nos testes iniciais foi o TrainingTime. Para diferentes grupos de dados, diferentes valores de trainingTime se mostravam mais eficientes. Por este motivo, foi decidido que tanto a temporalidade como o trainingTime seriam os parâmetros a serem variados neste trabalho. O trainingTime foi dividido em 9 configurações, de 500 vezes até 5000 vezes.

Neste trabalho foi utilizada a opção para treinamento Percentage Split com a configuração de treinamento/teste, sendo 90% para treinamento e o restante para a validação deste treinamento, caracterizando então um treinamento supervisionado. No Percentage Split o resultado é avaliado por uma porcentagem dos dados utilizados para treino, sendo uma parte para o treinamento e a parte restante para a validação da previsão, desse modo, é fornecido ao Weka apenas um conjunto de dados e o mesmo é dividido de maneira proporcional, sendo utilizada uma parcela para o treinamento e outra para a avaliação do algoritmo. A decisão do valor para treinamento e testes ser de 90% e 10% respectivamente foi tomada pelo fato de contermos um total de 780 instâncias. As 78 instâncias para testes representa uma boa quantidade, tendo em vista o escopo deste trabalho e que este número representa a quantidade de instâncias de um dia inteiro de coleta, caso esta coleta seja feita a cada 5 minutos.

5.5 DIVISÃO DAS SÉRIES TEMPORAIS

Como dito em uma sessão anterior, o algoritmo MultilayerPerceptron do Weka não trata a temporalidade dos dados. Então para este trabalho foi necessário que para cada instância de dados, fossem criados dados históricos à partir da instância anterior, gerando novas colunas na tabela de dados, como pode-se ver na Tabela 2.

A temporalidade destes dados é um dos parâmetros treinados e testados deste trabalho, ela foi dividida entre 7 configurações distintas. Estas configurações são:

- 15 minutos;
- 30 minutos;
- 45 minutos;
- 60 minutos;
- 75 minutos;
- 90 minutos;

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
C-30	H-30	L-30	O-30	V-30	C-25	H-25	L-25	O-25	V-25	C-20	H-20	L-20	O-20	V-20	C-15
252.42	253.2	251.64	253	139330	251.59	252.8	251.59	252.2801	129734	251.24	251.91	250.61	251.54	99882	251.42
251.59	252.8	251.59	252.2801	129734	251.24	251.91	250.61	251.54	99882	251.42	252.33	251.37	251.51	113014	250.53
251.24	251.91	250.61	251.54	99882	251.42	252.33	251.37	251.51	113014	250.53	251.65	250.41	251.55	114146	249.82
251.42	252.33	251.37	251.51	113014	250.53	251.65	250.41	251.55	114146	249.82	250.56	249.05	250.52	193769	250.8865
250.53	251.65	250.41	251.55	114146	249.82	250.56	249.05	250.52	193769	250.8865	251.06	249.67	249.79	106229	251.61
249.82	250.56	249.05	250.52	193769	250.8865	251.06	249.67	249.79	106229	251.61	252.16	250.7345	251.07	80981	251.2588
250.8865	251.06	249.67	249.79	106229	251.61	252.16	250.7345	251.07	80981	251.2588	252	251.14	251.63	55865	251.76
251.61	252.16	250.7345	251.07	80981	251.2588	252	251.14	251.63	55865	251.76	252.2	251.14	251.22	65941	251.47
251.2588	252	251.14	251.63	55865	251.76	252.2	251.14	251.22	65941	251.47	252.08	250.73	251.7126	71483	251.0741
251.76	252.2	251.14	251.22	65941	251.47	252.08	250.73	251.7126	71483	251.0741	251.66	251	251.47	47690	251.15
251.47	252.08	250.73	251.7126	71483	251.0741	251.66	251	251.47	47690	251.15	251.6	251.09	251.14	31750	250.34
251.0741	251.66	251	251.47	47690	251.15	251.6	251.09	251.14	31750	250.34	251.2536	250.33	251.2536	28514	250.96
251.15	251.6	251.09	251.14	31750	250.34	251.2536	250.33	251.2536	28514	250.96	251.07	250.34	250.4303	43893	251.03
250.34	251.2536	250.33	251.2536	28514	250.96	251.07	250.34	250.4303	43893	251.03	251.06	250.39	250.65	22176	250.675
250.96	251.07	250.34	250.4303	43893	251.03	251.06	250.39	250.65	22176	250.675	251.07	250.53	251.07	21489	250.18
251.03	251.06	250.39	250.65	22176	250.675	251.07	250.53	251.07	21489	250.18	250.98	250.164	250.73	34326	250.98
250.675	251.07	250.53	251.07	21489	250.18	250.98	250.164	250.73	34326	250.98	250.98	250.26	250.27	42874	250.9555
250.18	250.98	250.164	250.73	34326	250.98	250.98	250.26	250.27	42874	250.9555	251.44	250.9555	251	41401	251.12
250.98	250.98	250.26	250.27	42874	250.9555	251.44	250.9555	251	41401	251.12	251.25	250.535	250.92	18473	250.975

Tabela 2 – Tabela de dados divididos em 5 minutos.

- 120 minutos.

5.6 CONSIDERAÇÕES FINAIS

Neste capítulo foi apresentada a maneira que os dados foram coletados bem como a origem destes dados. Também foram definidas as ferramentas e metodologias utilizadas neste trabalho. Toda fundamentação teórica descrita nos capítulos anteriores contribui para o desenvolvimento deste capítulo. O próximo capítulo apresenta o desenvolvimento do trabalho e os resultados.

6 RESULTADOS OBTIDOS E DISCUSSÃO

Após a preparação dos dados e da escolha dos parâmetros a serem variados, deu-se início a próxima etapa deste trabalho, que é o treinamento e testes da RNA e testes. Como visto no capítulo anterior, foi usado o *Mean Absolute Error* para que sejam expressados os valores de erro dos testes.

Para efeito comparativo, para cada grupo de temporalidade, foi calculada a média da soma das variações, este calculo pode expressar quanto variava, em média, cada intervalo de cinco minutos de um determinado grupo de dados. Por exemplo, no grupo de testes em que a temporalidade se dá a cada 5 minutos no intervalo de 15 minutos, foi calculada a soma absoluta da variação entre 15min e 10min, somando com a variação de 10min com 5min e somando mais uma vez a variação de 5min com minuto 0, em seguida foi dividida essa soma por 3, calculando a média das variações. Desta forma foi possível criar um cálculo para comparar se o Mean absolute error calculado nos testes do Weka, está próximo desta média, assim como no cálculo manual, a Weka calcula a média da soma absoluta da variação entre o valor predito e o valor real.

Como visto no capítulo 5, os grupos de dados foram divididos entre dados somente com o valor do ativo financeiro, e o grupo de dados com os valores de abertura, fechamento, máximo, mínimo e volume. Na Tabela 3 pode-se analisar a saída dos testes para os dados que continham somente o valor do ativo financeiro da Tesla.

min/trainingTime	500	1000	1500	2000	2500	3000	3500	4000	5000	Média da variação
120	0.311	0.2882	0.2875	0.2886	0.2896	0.2909	0.292	0.2929	0.2942	0.2209
90	0.2957	0.2879	0.2798	0.2753	0.2734	0.2728	0.2723	0.2722	0.2723	0.252
75	0.3312	0.3347	0.3405	0.3416	0.3408	0.3394	0.338	0.3368	0.335	0.2638
60	0.4359	0.4245	0.4155	0.4102	0.4072	0.4055	0.4045	0.4039	0.4031	0.27
45	0.3805	0.3713	0.3777	0.3844	0.389	0.3922	0.3944	0.3959	0.3978	0.2738
30	0.3655	0.3422	0.3405	0.3438	0.3464	0.348	0.3493	0.3501	0.3508	0.2763
15	0.3634	0.3483	0.347	0.3426	0.3384	0.3351	0.3326	0.3307	0.3283	0.2795

Tabela 3 – Tabela dos resultados do grupo com o valor de fechamento.

Na Tabela 3 as colunas são representadas pelo TrainingTime, variando de 500 até 5000 vezes. Já as linhas são representadas pelo temporalidade, variando de 15 minutos até 120 minutos. O valor das células é mostrado em dólares. Como pode-se notar, quanto mais verde escuro a célula for, melhor o resultado, e menor o erro, e quanto mais vermelho forte, pior o resultado e maior o erro. Pode-se notar que mesmo na melhor configuração da RNA, o *Mean Absolute Error* foi de 0.2722 dólares, porém naquele grupo de 90 minutos, a média da soma das variações foi de 0.252. Pode-se aferir que em todas as configurações treinadas, nenhuma delas atingiu um *Mean Absolute Error* menor do que a média das variações.

O próximo grupo de dados testados foi o que continha os valores de abertura, fechamento, máximo, mínimo e volume. É muito difícil saber quais atributos uma RNA utiliza na hora de treinar o modelo e quais destes atributos melhoram ou pioram o seu desempenho. Por isso que foram realizados estes dois testes, para detecção e definição de qual se sairia melhor em cada treinamento. Na Tabela 4 pode-se analisar a saída dos testes realizados no grupo em que continham todos os atributos dos dados históricos.

min/trainingTime	500	1000	1500	2000	2500	3000	3500	4000	5000	Média da variação
120	0.258	0.2174	0.1902	0.1726	0.1625	0.1566	0.1528	0.1498	0.147	0.2209
90	0.2763	0.2515	0.2361	0.2236	0.2136	0.2049	0.1985	0.1935	0.1861	0.252
75	0.2108	0.1956	0.1869	0.1804	0.1773	0.1741	0.1713	0.169	0.1667	0.2638
60	0.223	0.198	0.1815	0.1691	0.1589	0.1514	0.1454	0.14	0.1326	0.27
45	0.2492	0.2229	0.2029	0.1879	0.1755	0.1647	0.1558	0.1496	0.1408	0.2738
30	0.2227	0.1812	0.1619	0.1494	0.1398	0.1335	0.1291	0.1255	0.1203	0.2763
15	0.2184	0.1831	0.1631	0.1507	0.1408	0.1333	0.1283	0.1254	0.1222	0.2795

Tabela 4 – Tabela dos resultados do grupo com o valor de todos atributos.

Na Tabela 4 também foram separados os trainingtime por colunas e a temporalidade em minutos por linhas. O valor das células é mostrado em dólares. Como no primeiro grupo, as células verde mais escuro são os dados com melhores resultados e menores erros e as células vermelho mais escuro são dados com piores resultados e maiores erros.

É possível inferir que neste grupo de dados com todos os atributos, os testes de modo geral se saíram melhores do que no grupo com somente os dados de fechamento. Ao compararmos os valores de Mean absolute error com a média da variação, nota-se que em muitos casos, o erro dos testes foi menor do que a média da variação.

Também é possível inferir que a melhor configuração se deu no grupo de treinamento que a temporalidade era de 30 minutos e o trainingtime era de 5000 vezes. Nesta melhor configuração, nota-se que o valor do erro é menor que a metade do valor da média de variação.

Como foi visto em capítulos anteriores, o foco principal deste trabalho é auxiliar um investidor doméstico na tomada de decisões na hora de investir na bolsa de valores, a fim de maximizar seu lucro. Os testes anteriores serviram para verificar quais seriam as melhores configurações para a RNA e se a configuração escolhida teria um erro menor do que a média da variação do seu respectivo grupo de temporalidade.

É importante ressaltar que durante este período de 01/02/2017 a 14/02/2017 em que os dados para treinamento foram coletados, o valor das ações da Tesla Motors subiu de \$252.42 dólares para \$280.98 dólares, isto representa uma alta de 11.3%. Desta forma, caso os dados estivessem em queda durante o período de coleta, não haveria como saber se os resultados dos testes da RNA seriam tão bons quanto os mostrados neste trabalho.

6.1 TESTES COM DADOS REAIS

Foram realizados testes com os dados reais do dia posterior a coleta dos dados para treinamento. Primeiramente foi escolhida a melhor configuração para o treinamento da RNA, como visto anteriormente, o trainingtime foi de 5000 vezes e a temporalidade foi de 30 minutos. Foram coletados os dados do dia 15/02/2017 e em seguida os dados foram transformados, de forma que eles trouxessem a temporalidade de 30 minutos bem como os dados de treinamento.

Como os dados do dia 15/02/2017 foram coletados no intervalo de 5 minutos, há um total de 78 instâncias para testes, porém na hora de popular os dados temporais, perdem-se 6 instâncias, pois para criar uma linha com dados de 30 minutos anteriores e divididos de 5 em 5, as 6 primeiras linhas da coleta são utilizadas para criar a sétima. A partir da sétima, se tem os valores dos 30 minutos anteriores.

Após o treinamento com a melhor configuração e os testes com os dados do dia posterior ao treinamento (15/02/2017), o Weka trouxe os valores reais, o valor predito e o erro (diferença entre real e o predito) como pode-se analisar na Tabela 5

Para poder calcular o possível lucro na compra e venda das ações, é necessário calcular a variação dos 5 minutos anteriores com o minuto real. Caso o valor seja positivo, é verificado uma valorização nas ações da empresa, caso negativo, este valor representa que o ativo financeiro desvalorizou. Sendo assim, foi adicionada na Tabela 5 uma coluna que continha o valor dos 5 minutos anteriores de cada instância, para que fosse possível calcular a variação real das ações. É possível também calcular se a RNA predisse valores condizentes com a realidade.

inst#	actual	predicted	error
1	280	279.803	-0.197
2	279.963	279.862	-0.101
3	279.84	279.978	0.138
4	280.08	280.022	-0.058
5	279.804	279.859	0.055
6	279.97	279.705	-0.265
7	280.03	280.049	0.019
8	279.854	279.898	0.044
9	280.3	280.171	-0.129
10	280.747	280.485	-0.262
11	280.16	280.338	0.178
12	280.11	280.206	0.096
13	280.548	280.328	-0.22
14	280.5	280.468	-0.032
15	280.51	280.718	0.208

Tabela 5 – Tabela dos resultados dos testes para o dia 15/02/2017.

Fonte: Autoria própria

Desta forma, também foram acrescentados uma coluna na qual é calculada a diferença entre os 5 minutos anteriores da instância e o valor que a RNA predisse daquela mesma instância, este valor calculado representa o quanto o valor do ativo financeiro irá valorizar ou desvalorizar segundo a RNA.

Na Tabela 6, foram adicionadas as colunas conforme mencionado anteriormente e feito os cálculos. Pode-se notar que quase sempre que a RNA predisse que o valor iria aumentar, ele realmente aumentou, e nos casos de desvalorização a RNA também previu isso. A coluna variaRNA representa a variação predita pela RNA, já a coluna variaReal representa a variação real da instância.

Para se ter uma noção de quanto lucro/prejuízo a previsão desta RNA poderia gerar, suponha-se que um investidor resolva comprar ações da Tesla Motors toda vez que a RNA preveja uma valorização no ativo financeiro de 30 centavos de dólar ou mais, portanto este investidor irá alimentar a RNA com os dados de testes para o dia em que deseja investir, e a partir dos resultados dos testes, ele compra determinada quantia de ativos toda vez que a RNA preveja um aumento de 30 centavos ou mais e nos próximos 5 minutos ele venda os ativos adquiridos anteriormente. Suponha-se também que este mesmo investidor realize esta ação ao decorrer do dia, nos horários de funcionamento da bolsa NASDAQ. Se este investidor tivesse

inst#	actual -5	actual	predicted	error	variaRNA	variaReal
1	279.744	280	279.803	-0.197	0.059	0.256
2	280	279.963	279.862	-0.101	-0.138	-0.037
3	279.9629	279.84	279.978	0.138	0.0151	-0.1229
4	279.84	280.08	280.022	-0.058	0.182	0.24
5	280.08	279.804	279.859	0.055	-0.221	-0.276
6	279.804	279.97	279.705	-0.265	-0.099	0.166
7	279.97	280.03	280.049	0.019	0.079	0.06
8	280.03	279.854	279.898	0.044	-0.132	-0.176
9	279.8545	280.3	280.171	-0.129	0.3165	0.4455
10	280.3	280.747	280.485	-0.262	0.185	0.447
11	280.7467	280.16	280.338	0.178	-0.4087	-0.5867
12	280.16	280.11	280.206	0.096	0.046	-0.05
13	280.11	280.548	280.328	-0.22	0.218	0.438
14	280.5476	280.5	280.468	-0.032	-0.0796	-0.0476
15	280.5	280.51	280.718	0.208	0.218	0.01

Tabela 6 – Tabela dos resultados dos testes para o dia 15/02/2017 e cálculos.

realizado estas ações no dia 15/02/2017, ao final do dia este mesmo investidor teria um lucro de \$10.3444 dólares para cada ativo comprado, como pode-se ver na Tabela 7 e no gráfico representado pela Figura ??.

É necessário ressaltar que este teste foi realizado em um dia que o valor do ativo financeiro da Tesla Motors teve alta de quase \$3 dólares ao fim do pregão, e que este lucro não pode ser comprovado em outras circunstâncias. Bem como não foi levado em considerações as taxas de operação das corretoras e nem de impostos sobre este valor movimentado. Foi realizado um caso de teste hipotético, porém com valores reais da cotação da empresa Tesla Motors.

inst#	actual -5	actual	predicted	error	variaRNA	variaReal
17	280.6	282.761	282.459	-0.302	1.859	2.161
18	282.761	283.135	283.855	0.72	1.094	0.374
21	283.46	284.67	284.538	-0.132	1.078	1.21
22	284.67	286.03	285.73	-0.3	1.06	1.36
39	282.81	283.615	283.461	-0.154	0.651	0.805
34	284.1754	284.66	284.685	0.025	0.5096	0.4846
20	282.9534	283.46	283.421	-0.039	0.4676	0.5066
40	283.615	284.408	284.074	-0.334	0.459	0.793
52	282.87	283.59	283.298	-0.292	0.428	0.72
57	283.6553	284.12	284.041	-0.079	0.3857	0.4647
50	282.84	283.15	283.206	0.056	0.366	0.31
45	283.59	284.3	283.919	-0.381	0.329	0.71
9	279.8545	280.3	280.171	-0.129	0.3165	0.4455
					Soma dos Lucros	10.3444

Tabela 7 – Tabela dos resultados e lucro no dia 15/02/2017.

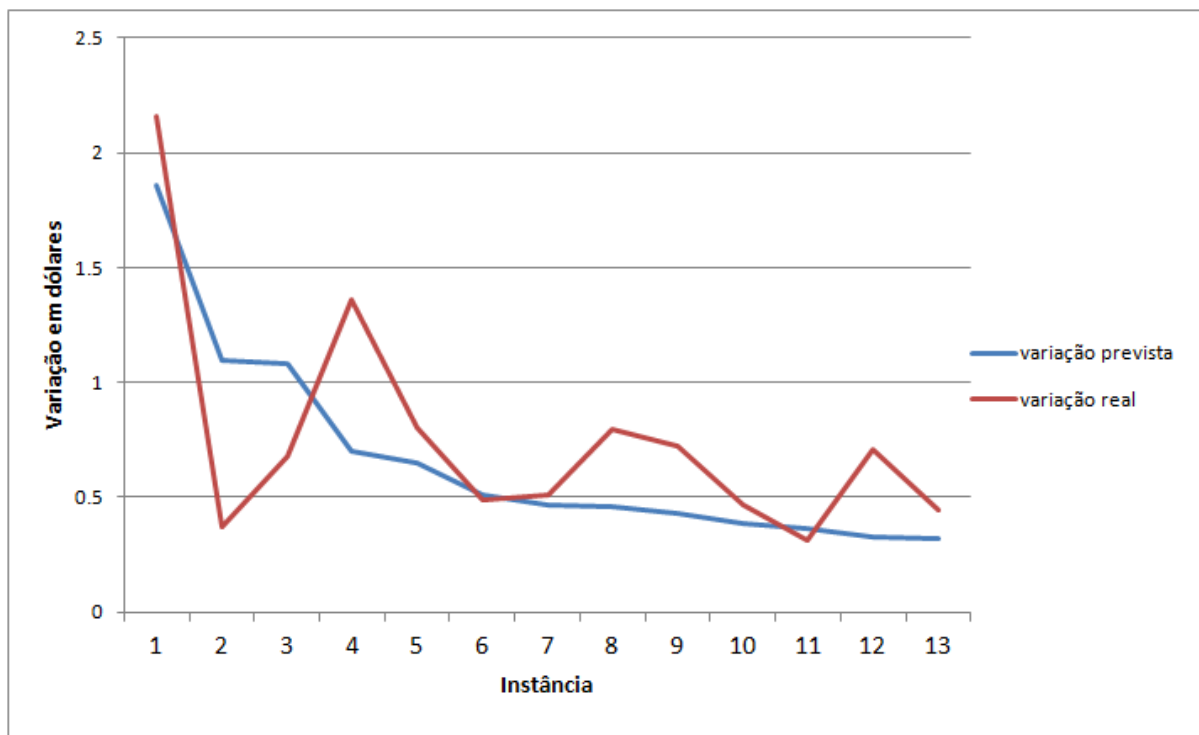


Figura 8 – Gráfico das variações previstas e reais

Fonte: Autoria própria

7 CONCLUSÃO

Neste trabalho foram apresentados e referenciados na fundamentação teórica os assuntos pertinentes a realização do mesmo. Esta fundamentação é de extrema importância para o entendimento do funcionamento dos mecanismos utilizados para a realização deste trabalho como um todo.

No capítulo Resultados Obtidos e Discussão, foram descritas os treinamentos e testes realizados no trabalho. A partir deles é possível inferir que o grupo com melhores resultados e menores erros foi o grupo de dados que continha os valores de todos os atributos coletados (abertura, fechamento, máximo, mínimo, volume). Algumas configurações para a RNA mostraram-se tão eficientes que na saída dos testes, o erro era metade do valor de variação. Porém não pode ser afirmado que este baixo valor de erro esteja diretamente relacionado a possíveis lucros, pois este erro é tomado da média de uma **soma absoluta**, sendo assim, existe a possibilidade de a RNA errar nas valorizações ou desvalorizações, fazendo com que o investidor tenha prejuízo em cada operação de compra e venda realizada, mesmo que o erro da RNA seja baixo.

Cabe ressaltar que todo o trabalho realizado objetiva auxiliar um investidor na hora de comprar e vender as ações de uma determinada empresa, baseando-se na cotação histórica dos valores destas mesmas ações. Este trabalho de forma alguma trata os fatores externos, como opinião pública sobre determinada empresa, notícias, especulações. Estes fatores podem influenciar muito no valor das ações. Porém um resultado mais preciso poderia ser obtido caso fosse realizado um estudo com maior número de dados. Possivelmente, uma análise com mais dias, bem como testes com mais dias, resultaria em uma maior precisão para as previsões, porém dado que o tempo e o escopo deste trabalho foram limitados, fazer tal análise foi inviável.

7.1 LIMITAÇÕES

No decorrer deste trabalho foram detectados as seguintes limitações:

Taxas de manutenção: Neste trabalho não foram levadas em considerações as taxas de manutenção e corretagem cobradas pelas corretoras. Estas taxas podem variar muito de acordo com a corretora de ações. A taxa de corretagem é cobrada por cada operação realizada, seja de compra ou de venda, podendo ser uma taxa fixa ou uma porcentagem do valor movimentado. Como demonstrado neste trabalho, para se maximizar o lucro é necessário realizar várias operações de compra e venda durante o dia de investimento, porém estas taxas podem ser um empecilho, visto que seriam cobradas diversas vezes.

Quantidade de dados: Pelo fato do tempo ser limitado para a realização deste trabalho, a quantidade de dados analisados também foi limitada.

Impostos e burocracia: Como neste trabalho, a bolsa de valores a ser investida é a NASDAQ, o investidor teria que ter uma conta em um banco internacional que operasse nos Estados Unidos, bem como teria que arcar com todas as despesas de impostos e taxas da receita federal. Estas taxas não foram calculadas neste trabalho, e podem desencorajar o investidor na hora de entrar no mercado de ações.

Falta de referencial: A falta de referencial relacionado a quantidade de parâmetros ou períodos para a aplicação de testes foi um fator limitante na hora de desenvolver o estudo.

7.2 TRABALHOS FUTUROS

Após o desenvolvimento deste trabalho, foram levantadas várias questões de interesse que permanecem em aberto, que poderiam dar segmento a este trabalho. Estas questões são:

- O estudo de mais empresas, relacionadas a empresa alvo, que podem influenciar no valor do ativo;
- Acrescentar atributos aos dados de treinamento, como coeficiente de correlação, médias móveis ou ainda desvio padrão, a afim de melhorar o desempenho da RNA;
- Reduzir ou aumentar o intervalo de tempo, seja de minuto a minuto, ou a cada hora;
- Aplicar formulas matemáticas e estatísticas aos dados antes dos mesmos serem

processados;

- Aplicar uma quantidade maior de dados ao estudo, não somente 10 dias.

REFERÊNCIAS

- AIKES, J. J. **Estudo da influência de diversas medidas de similaridade na previsão de séries temporais utilizando o algoritmo KNN-TSP**. Dissertação (Mestrado) — Unioeste, 2012.
- BATISTA, G. E. de A. P. A. **Pré-processamento de dados em aprendizado de maquina supervisionado**. Dissertação (Mestrado) — USP - São Carlos, 2003.
- BRAGA ANDRE PONCE DE LEON F. DE CARVALHO, T. B. L. Antonio de P. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: Grupo Editora Nacional, 2007.
- CARDON, D. N. M. A. **Introdução Às Redes Neurais Artificiais**. Dissertação (Mestrado) — UFRS, 1994.
- DAVIS, P. J. B. R. A. **Introduction to Time Series and Forecasting**. New York: Springer-Verlag, 2003.
- EBC, R. B. **Entenda como funciona o mercado de ações e o que são as bolsas de valores**. <http://www.ebc.com.br/noticias/economia/2015/08/entenda-como-funciona-o-mercado-de-acoes-e-o-que-sao-bolsas-de-valores>, 2015.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. de. **Inteligencia Artificial Uma abordagem de Aprendizado de maquina**. Rio de Janeiro: Grupo Editora Nacional, 2011.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. 1996.
- FIGUEREDO, C. J. **Previsão de Séries Temporais Utilizando a Metodologia box and Jenkins e Redes Neurais para Inicialização de Planejamento e Controle de Produção**. Dissertação (Mestrado) — UFPR, 2008.
- GEBHARDKIRCHGÄSSNER, J. **Introduction to Modern Time Series Analysis**. New York: Springer, 2013.
- HAYKIN, S. S. **Redes Neurais Princípios e Prática**. Rio de Janeiro: Bookman Editora, 2000.
- INFOMONEY, E. **Como Funciona o Mercado de Ações**. <http://www.infomoney.com.br/noticias/noticia/527576/como-funciona-mercado-acoes>, 2016.
- KOVÁCS, Z. L. **Redes Neurais Artificiais: Fundamentos e Aplicações**. [S.l.]: Collegium Cognitio, 1996.
- LEMOS, L. V.; LAGIOIA, U. C. T.; SANTOS, J. F.; LIBONATI, J. J.; Silva Filho, L. L. da. **Mercado de ações: Um estudo sobre o interesse do investidor em critérios não financeiros considerados na tomada de decisão**. 2013.
- MERENDA, L. **Previsões Marketing e gestão comercial**. [S.l.], 2014.
- MOBILIÁRIOS, C. D. V. **Mercado de valores mobiliários brasileiro**. Rio de Janeiro: [s.n.], 2014.

MORETTIN, C. M. d. C. T. P. A. **Modelos para previsão de séries temporais**. Rio de Janeiro: Instituto de matemática pura e aplicada, 1981.

NASDAQ. **NASDAQ**. <http://www.nasdaq.com/>, 2016.

O'NEIL, W. J. **How to make money in stocks**. New York: McGraw-Hill, Inc, 2009.

PYLE, D. **Data Preparation for Data Mining**. San Francisco - USA: Morgan Kaufmann Publishers, Inc., 1999.

REIS, M. M. **Análise de séries temporais**. [S.l.], 2007.

RUSSELL, P. N. S. **Inteligência Artificial**. Rio de Janeiro: Elsevier, 2010.

TUBINO, D. F. Manual de planejamento e controle de produção. 2006.

VELLASCO, M. M. B. R. Redes neurais artificiais. 2007.