

**UNIVERSIDADE FEDERAL DO MATO GROSSO DO
SUL**

**FACOM – FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA**

**PREDIÇÃO DO PREÇO DA ARROBA DO BOI
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

FELIPE FERRAZ DE SOUZA

ORIENTADORA: PROF^a. DRA. ANDREA TERESA RICCIO BARBOSA

CO-ORIENTADOR: PROF. DR. PEDRO PAULO PIRES

Campo Grande - MS
Dezembro/2021

**UNIVERSIDADE FEDERAL DO MATO GROSSO DO
SUL**

FACOM – FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**PREDIÇÃO DO PREÇO DA ARROBA DO BOI
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

FELIPE FERRAZ DE SOUZA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação Aplicada da Faculdade de Computação/FACOM - UFMS, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Campo Grande - MS
Dezembro/2021

AGRADECIMENTOS

Primeiramente gostaria de agradecer a minha orientadora Prof.^a Dr.^a. Andréa Tereza Riccio Barbosa, pela orientação, ensinamentos, paciência e incentivo durante a execução desse trabalho.

A Embrapa Gado de Corte e ao co-orientador Prof. Dr. Pedro Paulo Pires.

Aos professores da Faculdade de Computação da Universidade Federal do Mato Grosso do Sul que de alguma forma contribuíram durante todo o meu caminho.

Aos professores membro da banca pelas sugestões.

A todos os amigos que pude conviver durante a pós-graduação, pelo convívio e ajuda no decorrer do curso.

A minha família.

RESUMO

Nessa dissertação é proposto um modelo para predição do preço da arroba do boi gordo com referência no preço disponibilizado pelo CEPEA. Essa predição é feita com base na combinação de variáveis que interferem na composição do indicador, como por exemplo, o valor do dólar e o valor do milho. Após a identificação das variáveis que têm maior influência na composição do preço da arroba, para implementação do modelo, foi utilizada Redes Neurais Artificiais do tipo MLP (*Multilayer Perceptron*). Para realizar simulações com a rede MLP foi utilizado o *software* gratuito Weka (*Waikato Environment for Knowledge Analysis*) e várias configurações diferentes foram implementadas, tanto no número de variáveis, quanto na configuração do modelo de rede. Durante a execução dos testes, optou-se por adotar o modelo que apresentou o melhor resultado e com a configuração de rede possuindo o menor número de variáveis possível. A rede em alguns testes apresentou taxa de erro abaixo de 4%. Sendo assim, definiu-se a arquitetura, as variáveis de entrada e os parâmetros da rede mais indicada na predição do preço da arroba do boi para um futuro de noventa dias.

Palavras-chave: Predição, Redes Neurais Artificiais, Pecuária, Boi Gordo, Previsão de preço.

ABSTRACT

In this paper, a model for predicting the price of live cattle with reference to the price provided by CEPEA is proposed. This prediction is made based on the combination of variables that interfere in the composition of the indicator, such as the dollar value and the corn value. After identifying the variables that have the greatest influence on the composition of the price, to implement the model, Artificial Neural Networks of the MLP (Multilayer Perceptron) type were used. The simulation of the MLP network was used free software Weka (Waikato Environment for Knowledge Analysis) and several simulations were performed with different configurations, both in the number of variables and in the configuration of the network model. During the execution of the tests, it was decided to adopt the model that presented the best result and with the network configuration having the fewest possible variables. Some tests the network reached below 4% error rate. Therefore, the architecture and parameters of the network are more suitable for predicting the price of the cattle's.

Keywords: Prediction, Artificial neural networks, Livestock, Live Cattle, Price Forecast

LISTA DE FIGURAS

Figura 1 - Modelo não linear de um neurônio artificial (Neto, 2016).....	16
Figura 2 - Exemplo de arquiteturas, conexões e conectividades de RNAs (Braga <i>et. al.</i> , 2000)	18
Figura 3 - Exemplo de aprendizado supervisionado (Braga <i>et. al.</i> , 2000).....	20
Figura 4 - Exemplo de Rede Neural de Múltiplas Camadas (Haykin,2007).....	22
Figura 5 - Fluxo do algoritmo <i>backpropagation</i> (Braga <i>et. al.</i> , 2000).....	24
Figura 11 - Exemplo de um arquivo .arff	33
Figura 12 - Tela de configuração para MLP	35
Figura 10 - Fórmula de conversão de valores do preço a prazo para o VP da arroba de boi. (Adaptado CEPEA, 2019).....	39
Figura 13 - Evolução do preço da arroba do boi gordo	67
Figura 14 - Evolução do preço da arroba do boi em 2019	67
Figura 15 - Imagem Rede Neural	68
Figura 16 - Predição	69

LISTA DE TABELAS

Tabela 1 - Tabela adaptada com os resultados alcançados (Adisa <i>et. al.</i> , 2019)	31
Tabela 2 - Detalhamento dos Estudos Consultados	37
Tabela 3 - Variáveis Pesquisadas para serem consideradas como entrada da RNA e a fonte de onde podem ser obtidas.	44
Tabela 4 - Resultado dos primeiros testes iniciais	49
Tabela 5 - Resultado adicionando a variável valor_boi_australia	50
Tabela 6 - Resultado adicionando a variável valor_boi_eua ao restante das entradas.	51
Tabela 7 - Resultado comparativo com datas diferentes.....	52
Tabela 8 - Resultado comparativo adicionando as variáveis referente às chuvas das principais regiões do estado de São Paulo e Campo Grande.	53
Tabela 9 - Resultados da RNA considerando também a variável sazonalidade como característica de entrada.....	54
Tabela 10 - Resultado com a variável de sazonalidade e novos parâmetros de configuração.	55
Tabela 11 - Resultado do teste com a inserção das variáveis preço do bezerro e do frango. ..	56
Tabela 12 - Resultado do teste da RNA com a adição da variável IGPM.	56
Tabela 13 - Resultado teste com a Heurística “B” de configuração de rede.	58
Tabela 14 - Resultado inserindo a variável escala_abate	59
Tabela 15 - Resultado adicionando as variáveis valor_real_soja	60
Tabela 16 - Teste com a inclusão da variável	61
Tabela 17 - Tabela comparativa dos resultados da RNA entre períodos diferentes.....	62
Tabela 18 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2018).....	63
Tabela 19 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2018) incluindo as variáveis doença_animal e doença_humana	64
Tabela 20 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2019) incluindo as variáveis doença_animal e doença_humana	65
Tabela 21 - Resultado do teste com período o reduzido (01/01/2017 a 31/05/2021) incluindo as variáveis doença_animal e doença_humana	66
Tabela 22 - Resultado Predição dos 30 primeiros dias de predição para 90 dias futuros.	70
Tabela 23 - Resultado dos neurônios intermediários para o neurônio de saída da MLP	77
Tabela 24 - Valor dos Pesos das entradas para os neurônios intermediários da MLP	78

Tabela 25 - Valor dos Pesos da MLP.....79

Tabela 26 - Valor dos Pesos da MLP.....80

Tabela 27 - Valor dos Pesos da MLP.....81

LISTA DE ABREVIATURAS E SIGLAS

ABIEC - Associação Brasileira das Indústrias Exportadoras de Carnes

Adaline - *Adaptive Linear Element*

ARF - *Attribute-Relation File Format*

ARIMNA - *Autoregressive Integrated Moving Average*

BACEN - Banco Central

BM&F - Bolsa de Mercadorias e Futuro

BOVESPA - Bolsa de Valores de São Paulo

CDI - Certificado de Depósito Interbancário

CEPEA - Centro de Estudos Avançados em Economia Aplicada

CIIAGRO - Centro Integrado de Informações Agrometeorológicas

COVID-19 - *Corona Virus Disease*

CV - Coeficiente de Variação

EMATER/RS - Associação Riograndense de Empreendimentos de Assistência Técnica e Extensão Rural

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária

EQM - Erro Quadrático Médio

ESALQ - Escola Superior de Agricultura Luiz de Queiroz

FEALQ - Fundação de Estudos Agrários Luiz de Queiroz

FGV - Fundação Getúlio Vargas

GNU - *General Public License*

IBGE - Instituto Brasileiro de Geografia e Estatística

IBK - *Instance Based k*

IGP-M - Índice Geral de Preços do Mercado

INMET - Instituto Nacional de Meteorologia

IPC - Índice de Preço ao Consumidor

Land - Terra Cultivada (Land)

MCP - McCulloch-Pitts

MLP - *Multilayer Perceptron*

NTP - Número Total de Pesos

OMS - Organização Mundial da Saúde

PET - Evapotranspiração Potencial

PRE - Precipitação

RMSE - *Root mean squared error*

RNA - Rede Neural Artificial

RNAs - Redes Neurais Artificiais

SAD - Sistema de Apoio à Decisão

SARS-CoV-2 - *Severe Acute Respiratory Syndrome Coronavirus 2*

SM - Umidade do solo

SMOREG - *Sequential minimal optimization for Regression*

TMN - Temperatura Mínima (TMN);

TMX - Temperatura Máxima (TMX);

USDA - *United States Department of Agriculture*

WASDE - *World Agricultural Supply and Demand Estimates*

WEKA - *Waikato Environment for Knowledge Analysis*

SUMÁRIO

INTRODUÇÃO	8
1.1 Objetivo	10
1.2 Objetivos Específicos	10
1.3 Justificativa	11
1.4 Organização do trabalho	12
FUNDAMENTAÇÃO TEÓRICA	13
2.1 Produção do Gado de Corte	13
2.2 Redes Neurais Artificiais	15
2.2.1 Arquitetura	16
2.3 Aprendizagem	18
2.3.1 Supervisionado	19
2.3.2 Por Reforço	21
2.3.3 Não Supervisionado	21
2.4 <i>Multi Layer Perceptron</i>	21
2.4.1 Treinamento	22
2.4.2 Algoritmo de Treinamento (<i>Backpropagation</i>)	23
2.4.3 Desvantagens	24
TRABALHOS RELACIONADOS	26
3.1 Revisão Sistemática	26
3.2 Detalhamento dos trabalhos	27
3.3 Weka (<i>Waikato Environment for Knowledge Analysis</i>)	32
METODOLOGIA	38
RESULTADOS	42
5.1 Variáveis Pesquisadas e Coleta de Dados	42
5.2 Heurísticas Utilizadas na Implementação da Rede MLP	45
5.3 Implementação das RNAs	47
5.4 Teste Pré-Pandemia	47
5.4.1 Teste utilizando a Heurística A	52

5.4.2 Teste utilizando a Heurística B	57
5.5 Teste Pandemia.....	61
5.6 Predição	66
CONSIDERAÇÕES FINAIS	71
6.1 Conclusão	71
6.2 Trabalhos Futuros	73
REFERÊNCIAS.....	74
TABELA COM OS PESOS DA REDE MLP	77

Capítulo 1

INTRODUÇÃO

Em uma propriedade cuja atividade principal é a criação de gado, diversas variáveis influenciam na criação do rebanho. Dentre elas, a oferta de pasto e de alimentos complementares compatíveis com a atividade, a saúde do animal e também o clima. Este por sua vez, interfere diretamente sobre os aspectos anteriores, podendo comprometer os recursos financeiros do pecuarista, com uma eventual pressão sobre a produção de grãos e de gramíneas e, também, o bem-estar dos animais quando se deslocam a campo aberto.

Nota-se, que existem muitos aspectos que podem ser monitorados, praticamente em tempo real no interior de uma fazenda, mas segundo alguns especialistas do setor, é do lado de fora da propriedade que, normalmente, estão os maiores desafios desta atividade. Sendo alguns desses, o desempenho da criação de gado em países concorrentes, a disputa entre fornecedores por determinados mercados, a valorização da moeda, a relação entre oferta e demanda, entre outros (BeefReport, 2019).

Apesar dessa complexidade na produção do gado de corte, a pecuária brasileira é uma das mais eficientes do mundo. As extensões nacionais oferecem muitas possibilidades na seleção genética e na adoção de raças e cruzamentos que permitem um maior potencial de produção de proteína, proporcionando uma carne com as melhores características de sabor, suculência e benefício às necessidades energéticas do organismo humano. E devido a esses fatores que a carne bovina brasileira conquista muitos clientes, em diversos países no mundo (BeefReport, 2018).

Em 2017, oito estados foram responsáveis por 77,6% dos abates brasileiro. Mato Grosso do Sul e Mato Grosso juntos somaram em torno de 27% da produção mundial de carne bovina e abates legais, conforme dados do Instituto Brasileiro de Geografia e Estatística (IBGE, 2017). De acordo com a Associação Brasileira das Indústrias Exportadoras (ABIEC), em 2018, a

atividade da pecuária foi responsável por aproximadamente 597 bilhões de reais, representando 8,7% do PIB total do Brasil. Houve também o crescimento no volume da carne bovina produzida, com um total de 10,96 milhões de toneladas equivalente de carcaça. Desse total, 20,1% foi exportada e o restante destinado ao mercado interno (ABIEC, 2019).

O Centro de Estudos Avançados em Economia Aplicada (CEPEA) define que o boi gordo é considerado como indicador de preço, ou seja, bovino macho com 16 arrobas líquidas, ou mais de carcaça, e idade máxima de 42 meses. Este por sua vez influencia nos demais preços, como exemplo o boi magro. Schuntzemberger (2010) cita em seu trabalho que fatores temporais, tendências e variações irregulares como o preço de outras carnes, principalmente de frango, podem também exercer certa influência.

Devido, entretanto, às incertezas referentes aos preços que serão praticados na comercialização futura da produção, os produtores de gado de corte ficam submetidos a riscos, que podem comprometer os resultados dessa atividade. Partindo dessa premissa, é importante conhecer o comportamento das cotações da arroba do boi, para ter um diferencial mercadológico, para um melhor planejamento da produção de bovinos e para definição de estratégias de negociação (Correa *et. al.*, 2013).

Uma vez que o produtor decide fazer a engorda de seus animais sem saber o preço da arroba que receberá no futuro, sujeita-se ao risco do preço a ser praticado no momento da venda, não ser suficiente para cobrir os custos realizados no período da engorda. Evidencia-se que no processo de engorda, os preços podem sofrer variações de forma significativa, sob impacto de diversos fatores, que influenciam na oferta e na demanda do produto, tais como: variações no preço da carne suína ou frango; aumento dos preços dos insumos utilizados no processo de engorda, entre outros. Esses fatores podem ser suficientes para que o preço recebido, não cubra os custos de produção (Monteiro *et. al.*, 2010).

Para os produtores, portanto, é importante saber o preço que será praticado na venda para monitorar os gastos durante a engorda do animal, evitando assim, prejuízo financeiro. A utilização da Tecnologia da Informação (TI) pode ajudar a evitar esse problema e a maximizar os lucros, além de fornecer informações para monitoramento e controle.

Salienta-se ainda que as Redes Neurais Artificiais (RNAs), no âmbito da TI, possuem características essenciais para executar tarefas de predição e classificação, além da capacidade de aprender através de exemplos e generalizar as informações aprendidas. Existem muitas vantagens no uso de RNAs tais como: tolerância a falhas, capacidade de auto-organização e de processamento temporal (Braga *et. al.*, 2000). Tais vantagens, faz das RNAs uma ferramenta

computacional poderosa para solução de problemas complexos como, nesse caso, a predição da arroba de boi (Braga *et. al.*, 2000).

No cenário agropecuário, o uso de RNAs têm destaque no controle de índices zootécnicos com resposta ao comportamento animal, no consumo de energia metabólica, no manejo com base nas variáveis fisiológicas, no aspecto nutricional e produtivo, na sanidade do rebanho, no melhoramento genético, nas práticas de pré-abate, na determinação da qualidade do produto final, entre outros (Grassi, 2017).

Diante do exposto, este trabalho volta-se para a problemática da aplicação de técnicas computacionais na predição do valor da arroba do boi. Em particular, busca-se explorar e implantar o modelo *Muli Layer Perceptron* (MLP) e espera-se, como resultado, a validação do sistema para a predição do valor da arroba do boi a ser praticada no mercado financeiro.

1.1 Objetivo

Realizar previsão do preço da arroba do boi usando Rede Neural Artificial.

1.2 Objetivos Específicos

Os objetivos específicos são:

- definir como é realizada a precificação da arroba do boi no Brasil;
- determinar quais parâmetros influenciam na composição do preço da arroba do boi;
- verificar quais são as fontes possíveis de se obter os parâmetros determinados;
- compor os dados para treinamento da RNA do tipo *Multi Layer Perceptron* (MLP);
- realizar as simulações com a RNA do tipo MLP com sistema computacional;

- estabelecer a arquitetura adequada da rede MLP para generalizar a determinação da arroba do boi no Brasil;

1.3 Justificativa

O trabalho a ser desenvolvido foi sugerido, devido uma necessidade específica, pelos pesquisadores da Embrapa Gado de Corte de Campo Grande. Como citado anteriormente, o projeto tem por foco a predição ou previsão do valor da arroba bovina para o auxílio na tomada de decisão do produtor rural, pecuarista e demais envolvidos na produção de gado de corte, com a finalidade de evitar o desperdício de recursos, o aumento de produtividade e a renda.

Existem muitos trabalhos disponíveis na literatura que abordam o tema sobre predições para várias *commodities* e, alguns modelos estatísticos e computacionais foram utilizados na busca por soluções adequadas. Entretanto, por se tratar de um problema complexo, principalmente na identificação das variáveis relevantes para a composição do preço da arroba do boi, mais especificamente, existe a possibilidade de diversas pesquisas para a solução de tal problema. Optou-se, então, pelo uso da RNA do tipo MLP pela possibilidade de suas características apresentar resultados satisfatórios.

Quanto às contribuições do trabalho é possível identificar na área de computação, com o desenvolvimento computacional na solução da predição da arroba do boi, na área de pecuária, com possibilidade de obtenção de soluções práticas para os pesquisadores da EMBRAPA e, para a sociedade, mais especificamente aos pecuaristas que terão informações mais precisas nas realizações de transações comerciais, possibilitando a viabilização empresarial da atividade produtiva.

1.4 Organização do trabalho

A estrutura do trabalho está organizada da seguinte forma: o Capítulo 2 detalha os assuntos iniciais sobre o tema e sua Fundamentação Teórica. No Capítulo 3, apresenta-se alguns Trabalhos Relacionados, obtidos pela revisão sistemática realizada na literatura sobre a utilização de Redes Neurais Artificiais no âmbito da predição de preços da arroba do boi, entre outras *commodities*. No Capítulo 4, apresenta-se a metodologia considerando as etapas e os métodos realizados. O Capítulo 5 descreve os resultados e discussões do uso de Rede Neural Artificial na previsão do preço da arroba do boi. O último tópico, o Capítulo 6, apresentam-se os trabalhos futuros e as conclusões finais do trabalho.

Capítulo 2

FUNDAMENTAÇÃO TEÓRICA

2.1 Produção do Gado de Corte

Compreender o sistema de produção de gado de corte, mesmo que de uma maneira não tão aprofundada, é necessário para o entendimento sobre como é formado o preço do boi gordo e demais tipos de bovinos. Salientando-se que o valor de mercado se baseia no preço do boi gordo. Também pretende-se identificar, através das atividades econômicas da pecuária, quais são as principais variáveis que interferem na formação do preço da arroba do boi.

Conforme Cezar *et. al.* (2005) a produção de bovino tem como característica as fases de cria, recria e engorda. Tais fases podem ser atividades isoladas ou combinadas:

- cria: É composto pelo rebanho em produção, para o crescimento até a desmama para a comercialização. Esse processo dura entre 7 a 9 meses;
- cria e recria: A diferença para o anterior é que o macho é retido até 15 a 18 meses de idade;
- cria, recria e engorda: Os machos são vendidos como boi gordo para o abate com idade de 15 a 42 meses;
- recria e engorda: Inicia-se com o bezerro desmamado e termina com o boi gordo;

- engorda: É a fase de engorda do boi magro até atingir o peso para abate;

Devido à variedade e a diversidade socioeconômica das regiões em que os produtores se encontram, existem muitas formas diferentes da produção de carne bovina. Por isso, foi classificado e agrupado os sistemas de produção de acordo com os regimes alimentares dos rebanhos predominante no país, originando assim, três categorias de regime alimentar: Sistema Extensivo, Sistema Semi-intensivo e Sistema Intensivo (Cezar *et. al.*, 2005).

Os Sistemas Extensivos são os que utilizam pastagens nativas como fonte de alimento. Esse grupo possui uma alta variação de desempenho que é decorrente de fatores como o solo, o clima, a intensidade na utilização de pastagens, entre outros. O custo de produção é mais baixo, porém, as pastagens normalmente não possuem a quantidade necessária de minerais, sendo necessário uma suplementação de solo. No ano de 2005 este grupo representava 80% dos sistemas produtivos de carne bovina brasileira (Cezar *et. al.*, 2005), e 95% em 2020 conforme dados da EMBRAPA.

Os Sistemas Semi-intensivos também têm como base alimentar as pastagens e os suplementos minerais. Nesse modelo o objetivo é uma pecuária de ciclo mais curto, em que o animal recebe suplementos, em suas diversas fases de crescimento. As fontes energéticas mais utilizadas são milho, sorgo, aveia e milheto (Cezar *et. al.*, 2005).

Os Sistemas Intensivos têm foco na prática do confinamento na terminação do boi gordo. O confinamento começa antes do período de seca, em maio, e continua até depois do início do período chuvoso, em dezembro. O confinamento tem um período médio de 90 dias, no qual sua concentração ocorre de junho a outubro. Como fonte de alimentação é utilizado silagem de milho e sorgo (Cezar *et. al.*, 2005).

No estudo de Sachs e Pinatti (2007) verificou-se que os preços dos animais vivos possuem uma flutuação durante o ano, relacionado com disponibilidade de animais em ponto de abate, que é decorrente da maior ou menor disponibilidade de pastagens ao longo das estações do ano. Os autores também afirmam que o ciclo pecuário (variação do preço do gado e da carne, que acontece de tempo em tempo) deve ser levado em consideração, quando definidas as políticas de financiamento para o setor. Isso, porque, caso sejam adotadas em momento impróprio, podem reforçar as fases de preços altos, que prejudicam os consumidores, ou o de baixo preço, que prejudicam os produtores. Em se tratando de política para a pecuária, o adequado são as medidas que encurtem as fases ascendente ou descendente do ciclo pecuário ou que estabilizem o preço da carne em patamar mediano.

Mion (2019), identificou as principais variáveis que interferem na formação do preço do boi gordo no estado de São Paulo. As variáveis e a justificativa de cada uma são:

- preço do milho base Campinas – SP: devido a necessidade do estado de São Paulo importar 45% do milho que consome e pelo milho ser uma das principais fontes de energia no confinamento;
- preço do bezerro e do boi magro do Mato Grosso do Sul: ambos são utilizados para a reposição do rebanho;
- preço dos cortes bovinos base São Paulo: pelo fato da referência dos cortes (traseiro, dianteiro, ponta de agulha) ser a do estado de São Paulo;
- preço de carne *in natura* para exportação: boa parte da carne é para exportação, o que poderia gerar uma concorrência com relação ao destino da produção;
- preço do dólar: a cotação do dólar tem influência no preço final da carne exportada;
- histórico acumulado de chuva no estado de São Paulo: consiste em dados diários de chuvas acumuladas por 30 dias. Este volume de chuva tem influenciado o preço médio recebido pelo produtor;
- histórico de dias de escalas de abate: a capacidade diária de abate e a compra de boi gordo para compor sua escala, pode ser considerado uma variável que impacta diretamente no preço do boi gordo;

Além disso, Mion (2019) cita que além da atuação das forças de oferta e de demanda dentro da cadeia natural do setor pecuarista, condições externas também podem produzir novas forças de mercado e que influenciam no preço da arroba bovina. Neste último caso, o balanço entre oferta e demanda de carne no mercado, e também as variáveis dólar e precipitação, que requerem uma maior atenção. Maiores pesquisas serão realizadas nesse trabalho para verificar quais outros fatores podem influenciar no preço da arroba de boi.

2.2 Redes Neurais Artificiais

De um modo geral, as Redes Neurais Artificiais (RNAs) são um conjunto de técnicas computacionais, projetado para modelar a maneira como os organismos inteligentes adquirem conhecimento, por meio da experiência. Essa abordagem, consiste em solucionar problemas referentes à área da inteligência artificial, tendo fundamento em um sistema que simula o

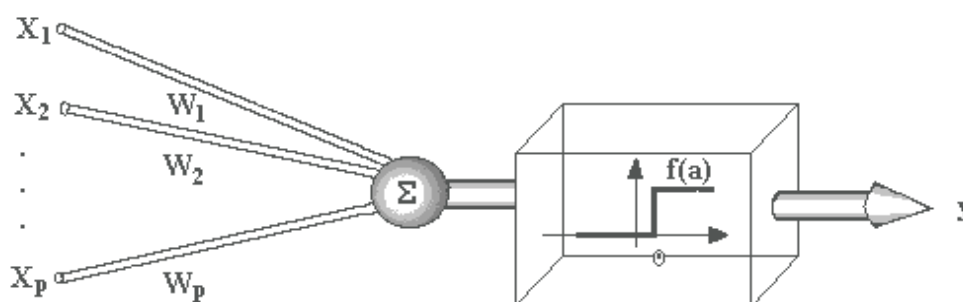
cérebro humano, até mesmo em seu comportamento, por meio de aprendizados, erros e descobertas (Cruz *et. al.*, 2018).

As RNAs são, frequentemente, utilizadas em resoluções de problemas complexos, no qual o comportamento das variáveis, não é rigorosamente conhecido. Normalmente, são utilizadas em diversas áreas de conhecimento, por exemplo, aplicações que envolvam reconhecimento de padrões, análise de séries temporais, diagnósticos médicos, previsão do mercado financeiro, entre outras. Uma de suas particularidades é aprender por meio de exemplo e de generalizar a informação aprendida, gerando um modelo não-linear, o que permite sua aplicação na análise espacial bastante eficiente (SPÖRL *et al.*, 2011).

2.2.1 Arquitetura

O modelo matemático proposto por Warren McCulloch e Walter Pitts foi inspirado na estrutura neural biológica de uma espécie inteligente. Denominado neurônio MCP (McCulloch-Pitts) é descrito por um conjunto de n entradas, as quais são multiplicadas por um determinado peso e, em seguida, os resultados são somados e comparados a um limiar (Braga *et. al.*, 2000). Na Figura 1 é representado um modelo artificial, onde x_1, x_2, x_p são as entradas da rede; w_1, w_2, w_n são os pesos, ou pesos sinápticos, associados a cada entrada; o Σ é o limiar de ativação e y é a saída do neurônio.

Figura 1 - Modelo não linear de um neurônio artificial (Neto, 2016)



De acordo com Braga *et. al.* (2000) a arquitetura de uma RNA é um parâmetro importante na sua concepção, de modo que ela restringe o tipo de problema que pode ser tratado pela rede. Salienta-se que uma camada única de nodos ou nós MCP, só conseguem resolver problemas linearmente separáveis. As redes recorrentes, que são estruturas de processamento

com a capacidade de representar uma grande variedade de comportamentos dinâmicos, são mais apropriadas para resolução de problemas no âmbito do processamento temporal.

A arquitetura é definida por Braga *et. al.* (2000):

- número de camadas da rede;
- número de nodos (ou nós) em cada camada;
- tipo de conexão entre os nodos;
- topologia da rede;

Na Figura 2 são apresentados alguns exemplos de arquiteturas de RNAs e suas variações com relação ao número de camadas, conexões e conectividade.

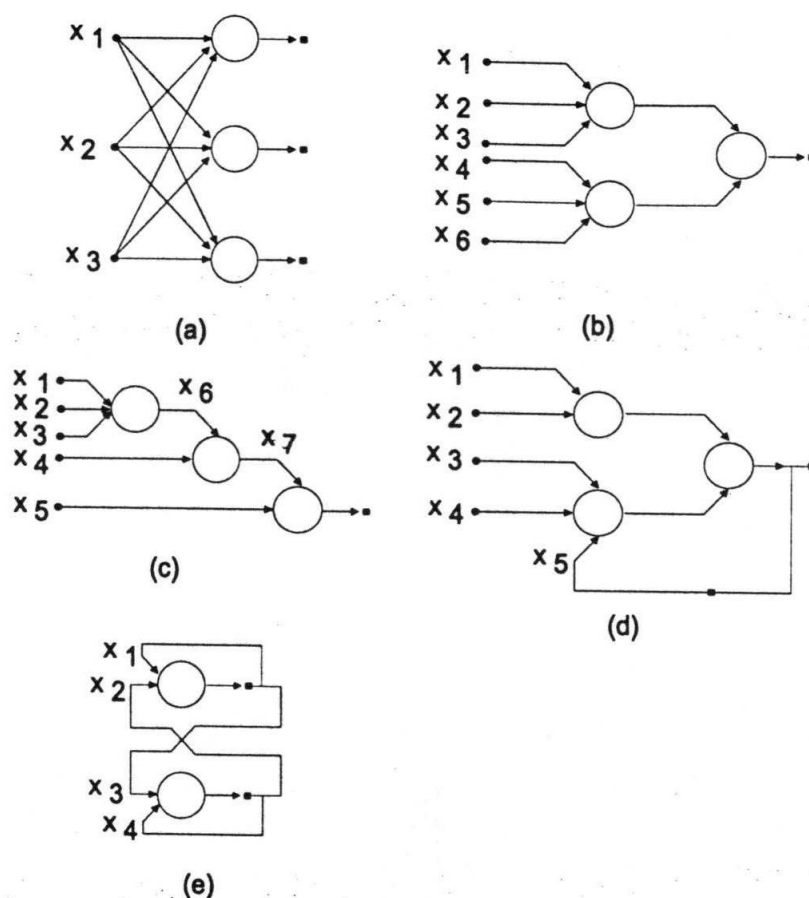
Com relação ao número de camadas, podem ser (Braga *et. al.*, 2000):

- redes de Camada Única: Possui apenas um nodo entre qualquer entrada e qualquer saída (Figuras a, e);
- redes de Múltiplas Camadas: Possui mais de um neurônio entre alguma entrada e alguma saída (Figuras b, c, d);

Os nodos podem ter conexões do tipo (Braga *et. al.*, 2000):

- *feedforward*: também conhecida como acíclica, a saída de um neurônio na *i*-ésima camada não pode ser usada como entrada de nodos em camadas de índice menor ou igual a *i* (Figura a, b, c);
- *feedback* ou cíclica: a saída de algum neurônio na *i*-ésima camada da rede é usada como entrada de nodos em camadas de índice menor a *i* (Figura d, e);

As RNAs podem ser classificadas também devido a sua conectividade. Pode ser rede fracamente ou parcialmente conectada (Figura b, c, d), ou completamente conectada conforme (Figura 2 a e F e).

Figura 2 - Exemplo de arquiteturas, conexões e conectividades de RNAs (Braga *et. al.*, 2000)

2.3 Aprendizagem

O fator de suma importância para uma rede neural é sua capacidade de aprender a partir de seu ambiente e de melhorar o seu desempenho através da aprendizagem. Uma RNA aprende sobre seu ambiente, por meio de um processo iterativo. De modo correto a rede se torna mais instruída sobre o seu ambiente, a cada iteração do processo de aprendizagem (Haykin, 2007).

Um conjunto de procedimentos bem definidos para a resolução de um problema de aprendizagem é chamado de algoritmo de aprendizagem. Não existe um único algoritmo de aprendizado, porém, há um conjunto de ferramentas representadas por diversos deles, sendo que a diferença ocorre pela maneira como o ajuste de pesos é realizado (Braga *et. al.*, 2000).

Existem diversos métodos para treinamento de redes e podem ser agrupados em dois paradigmas principais, aprendizado supervisionado e o não supervisionado. Outro método

conhecido é o aprendizado por reforço. A seguir será descrito sobre estes métodos com uma maior ênfase no método supervisionado, que será utilizado neste trabalho.

2.3.1 Supervisionado

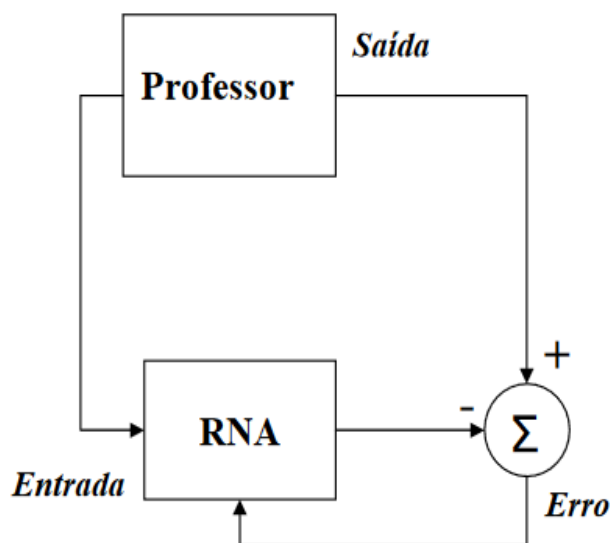
É considerado o método mais utilizado no treinamento das RNAs, devido sua facilidade em mensurar seus resultados. Possui esse nome porque a entrada e a saída desejada para a rede são fornecidas por um supervisor (chamado de supervisor ou professor).

O objetivo é ajustar os parâmetros da rede, de maneira a encontrar uma ligação entre a entrada e a saída fornecida. O supervisor indica um comportamento para rede (a relação entre a entrada e a saída), que possibilita direcionar o processo de treinamento.

No processo de aprendizagem, a rede tem sua saída atual (real) comparada com a saída desejada (fornecida pelo supervisor), obtendo-se a diferença entre esses valores, ou seja, o erro da resposta. A cada iteração, a rede faz uma comparação com a resposta desejada, fazendo um ajuste para minimizar o erro, desta forma, caminha para uma solução, caso exista uma solução possível. O *backpropagation* é o algoritmo mais conhecido para este método de treinamento (Braga *et. al.*, 2000). A Figura 3 ilustra o mecanismo do aprendizado supervisionado.

Uma desvantagem deste método é se não existir um supervisor, nesse caso, a rede não conseguirá aprender, pois não existe um exemplo da função a ser aprendida pela rede (Haykin, 2007).

Figura 3 - Exemplo de aprendizado supervisionado (Braga et. al., 2000)



O aprendizado supervisionado dispõe de duas maneiras de implementações. A *off-line*, no qual os dados do conjunto do treinamento não se alteram. E *on-line*, onde o conjunto de dados se altera de maneira contínua, e a rede deve permanecer em processo de adaptação. Os algoritmos mais conhecidos além do *backpropagation* são (Braga et. al., 2000):

- *perceptron*: primeiro modelo de rede neural proposto por Frank Rosenblatt em 1957 e utiliza o modelo do neurônio artificial de McCulloch & Pitts, funcionando como um classificador linear. O aprendizado é feito através da alteração dos valores dos pesos sinápticos entre os neurônios e utilizam algoritmos baseados na correção dos pesos, através do erro (Haykin, 2007);
- *Adaline (Adaptive Linear Element)*: Foi proposta por Widrow e Hoff em 1960 e apesar da mesma estrutura do *perceptron*, tem diferença no algoritmo de treinamento. A rede é adaptativa com método de aprendizado mais suave. Houve a inclusão de um algoritmo para o treinamento de redes de múltiplas camadas, a fim de realizar iterações para obter o ponto mínimo da função de erro, e desta maneira, fixar os valores dos pesos quando o mínimo for encontrado (Haykin, 2007);

2.3.2 Por Reforço

O aprendizado por reforço é um tipo de aprendizado supervisionado. A diferença é que a saída correta não é fornecida. A rede tem um *feedback* do ambiente, um sinal de reforço “sim” ou “não”, que é somente uma avaliação do desempenho da rede e não fornece informação de como melhorar o desempenho. Este método de aprendizado também pode ser chamado de aprendizado com crítica, ou seja, este crítico externo e responsável em avaliar a resposta fornecida pela rede (Braga *et. al.*, 2000).

2.3.3 Não Supervisionado

Como o próprio nome diz, não existe um supervisor ou professor para auxiliar no processo de aprendizado. Neste caso, somente as entradas estarão disponíveis. No momento que a rede determina a regularidade da entrada de dados, ela aprende pela experiência, classificando as características de entrada, e criando novas classes ou grupos. Para que isso se torne possível, é necessário a redundância nos dados de entrada. Sem redundância, não seria possível encontrar um padrão ou atributo nos dados de entrada (Braga *et. al.*, 2000).

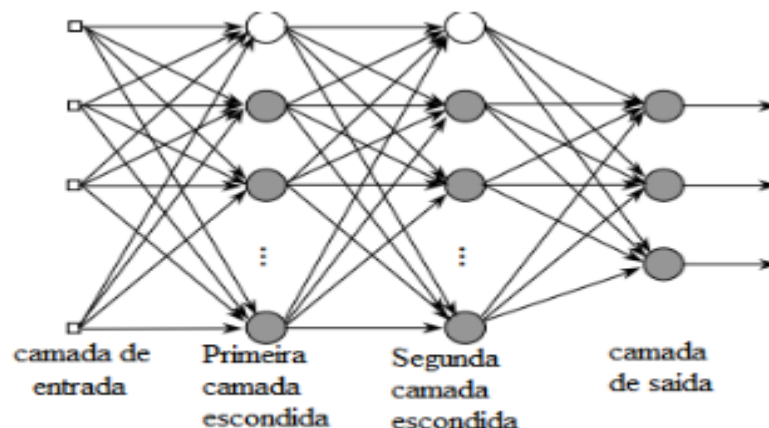
2.4 Multi Layer Perceptron

A Rede Neural do tipo *Multilayer Perceptron* (MLP) consiste em uma rede, cuja arquitetura, há a inclusão de camadas ocultas ou intermediárias, posicionadas entre a camada de entrada e a de saída da RNA. Essa arquitetura foi desenvolvida para atender as limitações da *perceptron* original e solucionar problemas, inclusive os não linearmente separáveis, e com regra de treinamento fundamentada no algoritmo *backpropagation* (Cruz *et. al.*, 2018).

A rede multicamadas é composta de camadas niveladas de neurônios. As entradas da rede estão na primeira camada. Esta camada distribui as informações da entrada para as camadas escondidas ou intermediárias, e na camada de saída é obtida a solução do problema. Poderão ter uma ou mais camadas escondidas, porém, na maioria das aplicações, considera-se apenas uma. Além disso os neurônios de uma camada estão conectados apenas aos neurônios

imediatamente posterior, não havendo realimentação, nem conexões entre neurônio da mesma camada e todas as camadas são totalmente conectadas, conforme Figura 4 (Santos *et al.*, 2005).

Figura 4 - Exemplo de Rede Neural de Múltiplas Camadas (Haykin,2007)



2.4.1 Treinamento

O treinamento supervisionado da MLP, conforme Cruz *et. al.* (2018), consiste em ajustar os pesos e os limiares dos neurônios, em cada camada, para que a saída seja a mais próxima do desejado. Existem duas dinâmicas de apresentação dos dados à rede, durante vários ciclos, em que o ajuste dos pesos vai sendo realizado (Braga *et. al.*, 2000):

1. Por padrão (*on-line* ou incremental): os pesos são reajustados após a apresentação de cada um dos n exemplos de treinamento, de maneira que no final de uma época (ou iteração) procede-se n ajustes. Nesta abordagem, a ordem da apresentação das observações é importante para o aprendizado da rede;
2. Por ciclo (*batch* ou *epoch*): os pesos são ajustados somente quando todos os exemplos de entrada e saída tiverem sido apresentados à rede, na fase de treinamento, de maneira que tenha apenas um ajuste, após uma época. Este treinamento é menos influenciado pela ordem de apresentação dos padrões.

Depois de selecionada a dinâmica do treinamento, é necessário determinar o momento final do treinamento, em que a rede é capaz de indicar valores apropriados para os pesos, a fim de que essa rede possa ser aplicada em predições, interpolações, entre outros.

Os critérios de parada do processo de treinamento são (Braga *et. al.*, 2000):

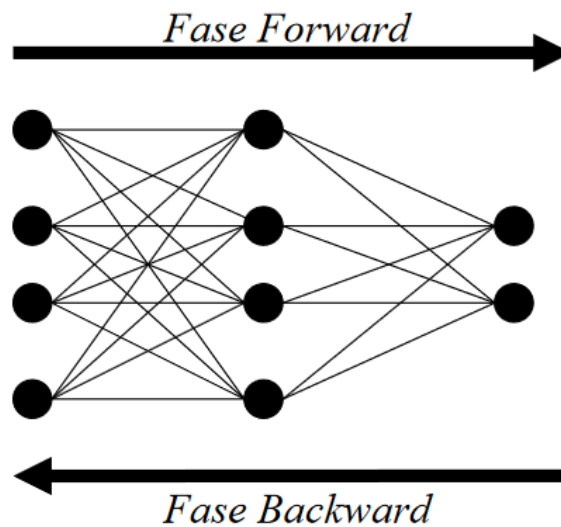
Número de ciclos (épocas ou iterações): número de vezes que o conjunto de treinamento é apresentado à rede. Deve-se dar atenção na quantidade de épocas a ser estabelecida, pois se

for um número excessivo pode levar a rede à perda de generalização. Caso seja um número pequeno de épocas, a rede pode não chegar ao seu melhor desempenho;

1. Estatística de custo: Podem-se usar valores de taxa de erro aparente, em estudos classificatórios ou coeficiente de determinação, em estudo de ajuste de modelo. Em geral, uma métrica utilizada é estabelecer um limite predefinido para o erro médio quadrático, para que a rede proporcione valores abaixo desse limite;
2. Combinação de métodos: Também é possível usar como método de parada a combinação dos métodos citados acima;
3. Validação: O treinamento é interrompido a cada conjunto de iteração, para que seja feito uma estimativa de erro da rede, sobre um novo conjunto de dados. Quando o erro medido no conjunto, independente dos dados, apresentarem crescimento, o treinamento realizado no conjunto original é encerrado. Nessa abordagem, o que se deseja é descobrir o momento em que a rede começa a perder generalização, confrontando o desempenho em fase de treinamento e de validação.

2.4.2 Algoritmo de Treinamento (*Backpropagation*)

O algoritmo *backpropagation* foi um dos principais responsáveis pelo ressurgimento do interesse em RNAs, devido a publicação de um trabalho *Parallel Distributed Processing* em 1986, apesar de ter sido proposto muito antes, com diferentes propósitos e pesquisadores. O *backpropagation* “funciona” como um mecanismo para correção dos pesos da rede. Através de um dado padrão de entrada é definida uma saída da rede, chamada de fase *forward*. Em seguida, com a saída desejada e a saída fornecida pela rede, é feita uma atualização do peso de suas conexões, chamada de fase *backward*. Na Figura 5 é apresentado o fluxo de processamento do algoritmo *backpropagation* (Braga et. al., 2000).

Figura 5 - Fluxo do algoritmo *backpropagation* (Braga *et. al.*, 2000)

Na fase *forward* a entrada é apresentada à primeira camada da rede. Nas camadas intermediárias é calculado o valor de saída, e serve como entrada para a camada seguinte. As saídas da última camada são comparadas as saídas desejadas. Na fase *backward* os nodos da camada atual ajustam seus pesos, a fim de reduzir seus erros. O erro das camadas intermediárias é calculado utilizando os erros dos nodos da camada seguinte, ponderado pelos pesos das conexões entre eles (Braga *et. al.*, 2000).

2.4.3 Desvantagens

Uma das dificuldades na utilização do algoritmo *backpropagation* é a disponibilidade de recursos computacionais que interfere na lentidão do algoritmo para superfícies mais complexas. Braga *et. al.* (2000) sugerem algumas técnicas para a acelerar o algoritmo *backpropagation*:

- utilizar taxa de aprendizado decrescente;
- adicionar nós intermediários;
- adicionar ruído aos dados;

Cruz *et. al* (2018) citam uma técnica proposta em 1986 por Rumelhart, chamada de *Momentum*, que é um valor constante usado para determinar o efeito das mudanças passadas dos pesos, na direção atual do movimento no espaço de pesos, e assim reduzir o tempo de treinamento e deixar a rede mais estável.

Existem várias modificações deste algoritmo para torná-lo mais eficiente. Na literatura como, por exemplo Haykin (2007), pode ser encontrado com maior detalhamento.

Capítulo 3

TRABALHOS RELACIONADOS

Este capítulo apresenta alguns trabalhos que podem contribuir com o desenvolvimento do tema desta dissertação, qual seja, predição do valor da arroba do boi, bem como a definição dos fatores relevantes na sua precificação.

3.1 Revisão Sistemática

A consulta foi realizada no período entre julho e dezembro de 2019, sendo que a data de publicação para busca dos trabalhos não foi delimitada, e foi feita a opção por artigos gratuitos ou que fossem acessíveis, dentro da universidade. Para a pesquisa foi utilizado o motor de busca *google.com*, a plataforma virtual da Biblioteca Brasileira de Teses e Dissertação – IBICT e o *software Harzing's Publish or Perish* que efetua busca de artigo em outros sites relacionados, tais como *Elsevier (Journal of Cleaner Production)*, *Emerald (International Journal of Sustainability in Higher Education)*, entre outros.

Na plataforma do IBICT foram encontradas 10 dissertações e 2 teses, relacionados ao tema. Desse resultado, após uma primeira análise, foram selecionados 6 trabalhos para uma leitura mais aprofundada. Verificou-se que nesta plataforma se obteve melhores resultados com palavras chaves na língua portuguesa.

Com o *software Harzing's Publish or Perish* foram encontrados cerca de 200 trabalhos. Desta forma, foi feito um filtro pelos trabalhos que possuíam maior relevância com o assunto

da pesquisa. Foram selecionados 7 artigos para uma análise mais criteriosa. Verificou-se que com este *software*, resultados mais relevantes foram obtidos quando se utilizou palavras chaves na língua inglesa. Também foi utilizado o *site* da Pós-Graduação da UFMS (<https://posgraduacao.ufms.br/portal/trabalhos>) para buscar trabalhos que possuíam relação com esta pesquisa. Foram selecionados 4 trabalhos desta base de dados.

As palavras chaves utilizadas nessas pesquisas foram:

- valor da arroba do gado;
- predição do valor da arroba do gado;
- preço da arroba do gado;
- predição + redes neurais + arroba de boi;
- *livestock market*;
- *price on cattle*;
- *feeder cattle*;
- *price discovery feeder cattle*

3.2 Detalhamento dos trabalhos

Sanders *et al* (2003) fizeram uma avaliação sobre as previsões de preço do gado, frango e porcos no *U.S Department of Agriculture* (USDA). Os autores identificaram que as previsões para o gado vivo poderiam ser melhoradas, caso utilizassem uma alternativa à série temporal. Mesmo assim o USDA identificou corretamente a direção do preço de suas previsões, em no mínimo 70%, no período amostrado.

O estudo examinou o desempenho das previsões trimestrais de preços do USDA, conforme relatado no *World Agricultural Supply and Demand Estimates* (WASDE). O padrão de comparação das previsões é gerado a partir de um modelo de série temporal univariada. Usando testes baseados em precisão, as previsões de preço do USDA produzem menor erro médio quadrático do que os gerados pelo modelo de série temporal, com a diferença, estatisticamente significativa, na previsão do desempenho.

Além disso, as previsões de preço do USDA são consideradas imparciais para gado e suínos, mas os preços dos frangos de corte são sistematicamente superestimados em 2,42%. Por

outro lado, o modelo de séries temporais subestima, constantemente, os preços em todos os três mercados.

As previsões do USDA são eficientes, pois não são conservadoras demais, nem muito extremas; no entanto, são ineficientes porque os erros de previsão tendem a ser repetidos, isto é, erros positivos são seguidos por erros positivos, principalmente no gado e mercados de frango.

Há evidências sugerindo que as previsões de gado não abrangem as previsões de séries temporais. Portanto, um previsor pode alcançar maior precisão, com menor erro médio quadrático, combinando as previsões de gado do USDA, com as de um simples modelo de série temporal.

De acordo com os resultados dessa análise, percebeu que a USDA fez um bom trabalho na previsão do preço dos animais num período trimestral. Mesmo assim, pode-se ainda melhorar a precisão das previsões para o preço do gado vivo, ou os intervalos de previsão podem ser aumentados, dinamicamente, para refletir mudanças de um mercado volátil.

Apesar de algumas deficiências, as previsões do USDA provavelmente agregam valor aos participantes do setor. Por exemplo, os profissionais podem usá-los para melhorar as previsões existentes. Mais importante, as previsões podem agregar valor aos participantes do mercado que não possuem a experiência, tempo ou recursos para gerar suas próprias previsões.

Medeiros *et. al.* (2006) fizeram através de uma abordagem metodológica, a utilização da modelagem *Autoregressive Integrated Moving Average* (ARIMA), na previsão do preço a ser recebido pela arroba do boi gordo.

A modelagem ARIMA consiste em um modelo ajustado, aos dados da série temporal, para melhor entender ou prever os padrões passados e futuros. Além de prever o preço a ser recebido, foi feita uma comparação do resultado obtido com o mercado real, disponível por meio da ESALQ/BM&F. Do ponto de vista teórico, o resultado do modelo apresentou coeficientes estatisticamente significantes, com indicadores dentro do intervalo de confiança previstos. Por exemplo, o intervalo de confiança em junho de 2005 vai de R\$ 50,33 a R\$ 53,36, tendo um valor previsto de 51,92. No ponto de vista prático, revelou-se um processo de previsão acurado, ou seja, 95,72% da variação do preço recebido pela arroba do boi gordo, apresentando resultados significativos e próximo aos valores praticados no mercado.

Ferreira *et.al.* (2011) utilizam RNAs para fazer a predição de preços, sobre o âmbito dos agronegócios. A base de dados foi da EMATER/RS, no período de 1992 a 2006, de quatro *commodities* específicas: soja, boi gordo, milho e trigo.

Os autores fizeram o uso do modelo de RNA do tipo MLP e os resultados mostraram a possibilidade do uso de Redes Neurais Artificiais, como uma alternativa da precificação dos preços das *commodities* em questão. Para cada uma dessas foi desenvolvida uma MLP, sendo que todas tinham a mesma arquitetura, de dois neurônios na camada de entrada, cinco na intermediária e um neurônio na saída.

Os dados utilizados na precificação do boi gordo foram o preço máximo e mínimo, em uma data específica. Os resultados obtidos podem ser considerados relevantes, devido às estatísticas R^2 , coeficiente de Theil, RMSE e MAPE, determinando que o modelo apresenta uma capacidade preditiva. Podendo assim, assegurar que a prática do uso de RNA pode ser agregada ao processo de tomada de decisão, com relação aos preços de *commodities* agrícolas. Tais resultados podem demonstrar um ponto de partida, para a melhoria aos gestores rurais.

Aguiar (2016) faz uma análise sobre quais variáveis afetam o preço do boi gordo, no estado de São Paulo. Foi observado que no período, entre 2003 e dezembro de 2015, a cotação do preço do boi gordo foi influenciada pela renda nacional e internacional de importantes países importadores, tendo como fator o preço do barril do petróleo, a taxa de câmbio e o índice pluviométrico das principais regiões produtoras do estado, sendo esses dados coletados diretamente do Centro Integrado de Informações Agrometeorológicas (CIIAGRO).

No que diz respeito a renda nacional, 14,35% das variações foram transmitidas para cotação do preço do boi gordo, no mesmo período, e 11,84% acontecem com defasagem de um mês. Esse ajuste de preço imediato, deve-se ao curto processo produtivo entre o abate do animal e a disponibilidade para o consumo. Desta forma, os frigoríficos podem ajustar o preço de compra, para se adaptar a variações na demanda.

Com relação às variáveis que se referem ao mercado internacional, observa-se que os ajustes de longo prazo são mais significativos, que os de curto. Essa diferença temporal deve-se aos contratos comerciais, que como são com prazos definidos, não se ajustam em velocidade compatível com as variações na demanda.

A taxa de câmbio, por exemplo, apresentou efeitos de 5,24% no curto prazo e 24,49% no longo prazo. A variável do índice de chuva apresentou 18,37% das variações que são transmitidas para a cotação do preço do boi gordo, com defasagem temporal de cinco meses. O objetivo desse estudo foi fornecer informações para que os pecuaristas possam tomar decisões, de acordo com suas necessidades de compra ou venda de gado.

Grassi (2017) utilizou RNA para previsão do pré-parto de vacas da raça Holandesas, possibilitando o acompanhamento do parto pelo pecuarista, com o intuito de evitar a incidência de natimortos, ou seja, bezerros que nascem mortos ou morrem, pouco após o nascimento. Para

a predição do parto do gado Holandês foi utilizado o modelo *Multilayer Perceptron* (MLP) e para a simulação da rede foi utilizado o *software* Weka (*Waikato Environment for Knowledge Analysis*). Desta forma, várias simulações foram realizadas. As simulações cuja a rede teve como saída a hora restante para a ocorrência do parto, não obteve resultados satisfatórios. Novos testes foram realizados tendo como saída da rede MPL, o momento de parto e não parto ou momento de parto, pré-parto e não-parto. Essas configurações foram as que apresentaram melhores resultados. Para o teste que apresentou o melhor desempenho, foi coletado a temperatura do retículo-ruminal do animal e utilizada 51 vacas, totalizando 153 exemplos. Nesse teste foi utilizado um filtro de média móvel para garantir que as temperaturas estivessem suavizadas ou sem ruídos.

O exemplo de teste apresentado chegou a 100% de taxa de acerto e devido a esses fatores foi concluído que esse modelo é indicado para a automatização do processo de predição do parto em vaca da raça Holandesa.

Conclui-se com os resultados alcançados, que o desempenho não está relacionado, necessariamente, com a complexidade do modelo adotado. Também foi observado que, a classificação dos modelos varia de acordo com o horizonte de previsão e a medida de erro adotada. Um único modelo de RNA superou os demais, para horizonte mais longos, e a combinação de previsões, mesmo que simples, pode trazer melhorias na acurácia.

Os resultados apresentados nesse trabalho, portanto, pode contribuir com os pesquisadores com foco em produção de boi gordo, auxiliando na elaboração e uso de previsões para a *commodity* em questão. Contudo há muito ainda a se explorar, visto que são diversas variáveis que podem compor o preço da arroba bovina e que não foram abordadas.

Adisa *et. al.* (2019) utilizaram RNAs para prever a produção de milho nas principais províncias da África do Sul. As análises de previsão foram realizadas usando as seguintes variáveis climáticas:

- precipitação (PRE);
- temperatura máxima (TMX);
- temperatura mínima (TMN);
- evapotranspiração potencial (PET);
- umidade do solo (SM);
- terra cultivada (Land) para milho;

Os conjuntos de dados analisados foram de 1990 a 2017 e foram divididos em dois segmentos, com 80% usados no treinamento do modelo e os 20% restantes nos testes. Os resultados indicaram que PET, PRE, TMN, TMX, Land e SM, com duas camadas ocultas seriam a melhor arquitetura para prever a produção de milho na província de *Free State*, enquanto a TMN, TMX, PET, PRE, SM e Land, também com duas camadas ocultas, sendo uma com 7 neurônios e a outra com 8 neurônios, foram a melhor combinação, para prever milho na província de *KwaZulu-Natal*.

Além disso, o TMN, SM e Land e o TMN, TMX, SM e Land, com duas camadas ocultas, sendo uma com três e a outra com quatro neurônios, foi a melhor arquitetura para a previsão de milho nas províncias de *North West* e *Mpumalanga*, respectivamente.

A comparação entre a produção real e a prevista de milho, usando os dados do teste indicaram precisão de desempenho de 0,75 para *Free State*, 0,67 para *North West*, 0,86 para *Mpumalanga* e 0,82 para *KwaZulu-Natal*. Além disso, um declínio na projeção da produção de milho foi observado em todas as províncias selecionadas (exceto na província de *Free State*), de 2018 a 2019. Assim, o modelo desenvolvido pode ajudar a aprimorar o processo de tomada de decisão por parte dos agricultores. A Tabela 1 apresenta os principais resultados obtidos com este estudo.

Tabela 1 - Tabela adaptada com os resultados alcançados (Adisa et. al., 2019)

Província	Combinação de Variáveis	Neurônio Oculto	Escore
KwaZulu-Natal	PET, PRE, TMN, TMX, Land, SM	7,8	93,90%
North West	TMN, Land, SM	3,4	73,74%

Neto (2016) propôs o desenvolvimento de um *software* com conceito de Sistema de Apoio à Decisão (SAD), visando o controle do confinamento do gado, a partir das informações obtidas de dados oriundos de uma balança de passagem, dos valores nutricionais dos alimentos consumidos pelo animal e do mercado bovino. Dentre as funcionalidades a de maior relevância para a tema do trabalho em desenvolvimento atual, destaca-se a implementação de uma RNA para a predição do valor da arroba do boi. Foram utilizados três algoritmos: MLP, SMOREG e IBK. O teste resumia-se em reunir índices satisfatório da predição da arroba bovina, para um confinamento de cerca de 100 dias. O *software* utilizado para realizar as simulações das predições foi o Weka (*Waikato Environment for Knowledge Analysis*). Para o conjunto de

dados que compuseram as variáveis, foram utilizados históricos do valor da arroba bovina entre os anos de 2007 e 2015. A menor taxa de erro utilizada com o algoritmo MLP foi de 10%.

Ao finalizar os testes foi observado que os resultados obtidos não foram satisfatórios, devido aos erros estarem acima da margem de tolerância, gerando um erro de precisão, quando as análises de custos fossem realizadas.

3.3 Weka (*Waikato Environment for Knowledge Analysis*)

Muitas ferramentas que simulam uma RNA estão disponíveis no mercado, tanto comercial como as de código aberto, entretanto, a implementação da rede depende do conhecimento de como o *software* deve ser implementado e das particularidades e condições inerentes do problema a ser resolvido. Os principais motivos para a escolha do Weka foram (Amaral, 2016):

- pode ser facilmente baixado sem custo de aquisição;
- possui um conjunto grande de algoritmos capaz de executar as mais diferentes tarefas de aprendizado de máquina, dentre eles o *MultiLayer Perceptron*;
- é uma ferramenta consolidada no mercado;
- possui uma interface gráfica disponível tornando o aprendizado mais intuitivo.

O Weka possui um formato próprio de arquivo para o aprendizado de máquina. Possui a extensão *.arff* (*Attribute-Relation File Format*) que é um arquivo texto no padrão ASCII e pode ser aberto com qualquer editor de texto. Contém um cabeçalho (*@relation*) com as informações e tipos de variáveis do arquivo e uma seção com os dados (*@data*) utilizados na aprendizagem. No pacote de instalação do Weka estão disponíveis cerca de 20 arquivos de exemplos que ajudam no estudo da ferramenta.

Figura 6 - Exemplo de um arquivo .arff

```
@relation weather

@attribute outlook {sunny, overcast, rainy}
@attribute temperature numeric
@attribute humidity numeric
@attribute windy {TRUE, FALSE}
@attribute play {yes, no}

@data
sunny,85,85,FALSE,no
sunny,80,90,TRUE,no
overcast,83,86,FALSE,yes
rainy,70,96,FALSE,yes
rainy,68,80,FALSE,yes
rainy,65,70,TRUE,no
overcast,64,65,TRUE,yes
sunny,72,95,FALSE,no
sunny,69,70,FALSE,yes
rainy,75,80,FALSE,yes
sunny,75,70,TRUE,yes
overcast,72,90,TRUE,yes
overcast,81,75,FALSE,yes
rainy,71,91,TRUE,no
```

Figura 12 mostra a tela de configuração da RNA do tipo MLP e os principais itens de configuração são:

- **GUI:** mostra uma interface com uma visualização gráfica da rede e pode ser usada para criação e alteração;
- **autoBuild:** Quando “setado” em *true*, as camadas ocultas são adicionadas e conectadas. Há uma camada oculta por padrão, mas se não houver *autoBuild*, ela não será exibida e não haverá conexão;
- **batchSize:** O número preferencial de instâncias do vetor mantidas no *cache* uma vez, se a predição em lote estiver sendo executada. Instâncias podem ser fornecidas, mas isso dá às implementações uma oportunidade de especificar um tamanho de lote preferido;
- **debug:** pode enviar informações adicionais ao console se “setado” como *true*;
- **decay:** fará com que a taxa de aprendizagem diminua. Dividindo a taxa de aprendizado inicial pelo número da época, para determinar qual deve ser a taxa de aprendizado atual. Isso pode ajudar a impedir a rede de divergir da saída de destino, bem como melhorar o desempenho geral;
- **doNotCheckCapabilities:** Quando definido, os recursos do classificador não são verificados antes da criação do classificador (pode reduzir o tempo de execução);
- **hiddenLayers:** configuração da quantidade de nós e quantidade de camadas da camada oculta da rede;

- *learningRate*: taxa de aprendizagem da rede;
- *Momentum*: refere-se a atualização de peso das conexões do nós;
- *nominalToBinaryFilter*: representa o filtro que irá pré-processar as instâncias. Isso pode ajudar a melhorar o desempenho se houver atributos nominais no conjunto de dados;
- *normalizeAttributes*: Se “setado” como *true*, normalizará os atributos, podendo ajudar a melhorar o desempenho da rede;
- *normalizeNumericClass*: Normalizará a classe se for numérica. Pode ajudar a melhorar o desempenho da rede, normalizando a classe para o intervalo $[-1,1]$. É feito apenas internamente, a saída será reduzida para a faixa original;
- *numDecimalPlaces*: Número de casas decimais usadas para a saída no modelo;
- *reset*: Permite que a rede seja redefinida com uma taxa de aprendizado mais baixa. Se a rede divergir da resposta, será redefinido automaticamente a rede com uma taxa de aprendizado mais baixa e começará o treinamento novamente. Esta opção estará disponível somente se a GUI estiver “setada” como *false*;
- *resume*: Define se o classificador pode continuar o treinamento após realizar o número solicitado de iterações.
- *seed*: usado para inicializar o gerador de números aleatórios. Números aleatórios são usados para definir os pesos iniciais das conexões entre os nós e para embaralhar os dados de treinamento.
- *trainingTime*: número de épocas para treinamento;
- *validationSetSize*: é o percentual do conjunto de validação. O treinamento continuará até que o erro no conjunto de validação tenha piorado consistentemente ou até que o tempo de treinamento seja atingido.
- *validationThreshold*: Usado para encerrar o processo de aprendizagem. O valor indica quantas vezes consecutivas o erro do conjunto de validação pode piorar antes do término do treinamento.

Figura 7 - Tela de configuração para MLP

The screenshot shows the 'weka.gui.GenericObjectEditor' window for the 'weka.classifiers.functions.MultilayerPerceptron' classifier. The 'About' tab is active, displaying a description: 'A classifier that uses backpropagation to learn a multi-layer perceptron to classify instances.' Below this, there are two buttons: 'More' and 'Capabilities'. The main configuration area contains various settings, each with a label and a value field or dropdown menu. At the bottom, there are four buttons: 'Open...', 'Save...', 'OK', and 'Cancel'.

Property	Value
GUI	False
autoBuild	True
batchSize	100
debug	False
decay	False
doNotCheckCapabilities	False
hiddenLayers	a
learningRate	0.3
momentum	0.2
nominalToBinaryFilter	True
normalizeAttributes	True
normalizeNumericClass	True
numDecimalPlaces	2
reset	True
resume	False
seed	0
trainingTime	500
validationSetSize	0
validationThreshold	20

O Weka possui um item *Summary* no qual é possível visualizar as métricas de avaliação do modelo. As métricas por padrão quando utilizado a MLP são:

- *Correlation coefficient*: Indica o quanto o conjunto de dados está relacionado. Quanto mais próximo de 1 ou -1 pode-se dizer que existe uma alta correlação. Se for uma correlação positiva, indica que os valores aumentam juntos e negativa indica que quando um aumenta o outro diminui.
- *Mean absolute error*: É a média da diferença absoluta dos valores previstos e valores observados. Quanto menor esse valor, melhor é o desempenho do modelo.
- *Root mean squared error (RMSE)*: Nessa métricas os erros são elevados ao quadrado antes de serem calculados. Basicamente implica que o RMSE atribui

um peso maior a erros maiores, indicando que o RMSE é muito mais útil quando grandes erros estão presentes e afetam drasticamente o desempenho do modelo. Assim como na *Mean Absolute Error*, quanto menor o valor, melhor é o desempenho do modelo.

- *Relative absolute error*: O erro relativo é calculado com o erro absoluto dividido pelo tamanho do valor exato. O erro absoluto é o tamanho da diferença entre o valor real e o valor estimado. Este resultado é mostrado em percentual.
- *Root relative squared error*: É a média dos valores reais. Assim, o erro quadrático relativo considera erro quadrático total e o normaliza dividindo pelo erro quadrático total.

Total Number of Instances: Número de instancias utilizadas para avaliação.

Na Tabela 2 pode ser observado o detalhamento dos artigos.

Tabela 2 - Detalhamento dos Estudos Consultados

Autor	País	Objetivo	Variável Independente	Modelos	Resultados
Sanders <i>et al.</i> (2003)	Estados Unidos	Avaliação sobre as previsões de preços da USDA; Previsão dos preços num período trimestral.	Serie de preços de boi, frango, porco	Comparação feita a partir de um modelo temporal univariada.	Previsão da USDA eficiente, porem pode ser melhorada para o preço do gado. Identifica cerca de 70% as variações dos preços.
Medeiros <i>et. al.</i> (2006)	Brasil	Previsão utilizando o fechamento mensal entre jun/2005 a jun/2006	Preço a ser recebido pela arroba do boi gordo	ARINA	Processo de previsão mostrou-se acurado com 95,72%.
Ferreira <i>et.al.</i> (2011)	Brasil	Previsão de preço médio mensal no âmbito do agronegócio, no período de 1992 a 2006.	Preço a ser recebido pela soja, boi gordo, milho e trigo.	Redes Neurais Artificias do tipo <i>Multilayer Perceptron</i>	Os resultados obtidos foram considerados relevantes em que o erro médio percentual absoluto ficou em torno de 5,73%
Aguiar (2016)	Brasil	Análise de quais variáveis afetam o preço do boi gordo no período entre 2003 a 2015.	Além da renda nacional e internacional, foi observado também o preço do barril de petróleo, taxa do câmbio e índice pluviométrico.	-	A curto prazo, foi observado que a renda nacional tem mais influência no preço. Já a longo prazo, são as variáveis que se referem ao mercado internacional.
Grassi (2017)	Brasil	Realizar a predição do momento do parto das vacas Holandesas	Verificar o momento do parto, pré-parto e não parto conforme a temperatura do animal.	Redes Neurais Artificias do tipo <i>Multilayer Perceptron</i>	O modelo desenvolvido chegou a 100% de acerto, com um teste de 153 exemplos
Adisa <i>et. al.</i> (2019)	África do Sul	Prever a produção do milho conforme o clima das regiões da África do Sul.	O conjunto de dados pluviométricos foi coletado entre 1990 e 2017.	Redes Neurais Artificias do tipo <i>Multilayer Perceptron</i>	O modelo teve resultados satisfatórios quando comparado a previsão com a produção real do milho. Os exemplos mostram a acurácia registrada de até 90%.
Neto (2016)	Brasil	Prever o preço da arroba do boi gordo.	Obter índice num período de 100 dias. Os dados foram coletados entre 2007 a 2015	Redes Neurais Artificias do tipo <i>Multilayer Perceptron</i> ; IBK; SMOREG;	Os resultados obtidos não foram satisfatórios conforme o autor, indicando que esse assunto precisa ser mais explorado.

Capítulo 4

METODOLOGIA

Com o objetivo de fazer a predição do preço da arroba do boi, usando Rede Neural Artificial do tipo *Multi Layer Perceptron*, algumas etapas foram desenvolvidas, principalmente, para definição e obtenção dos dados que fizeram parte do banco de dados, utilizados no treinamento da rede.

As etapas são descritas a seguir:

Etapa 1: Definir como é realizada a precificação da arroba do boi no Brasil

Nota-se que os preços da arroba do boi são formados a partir da associação entre oferta e demanda do produto e, o boi gordo é considerado como o indicador do valor. Normalmente, esses indicadores utilizados são disponibilizados pelo convênio entre a Fundação de Estudos Agrários Luiz de Queiroz (FEALQ) e a Bolsa de Valores, Mercadorias e Futuros de São Paulo BM&FBOVESPA (Pereira, 2017).

O indicador é feito através de uma média diária dos valores, praticados nas principais regiões produtoras do estado de São Paulo. Os pecuaristas, comerciantes, representantes de frigoríficos, escritórios de compra e venda de gado e leiloeiras são definidos como os agentes colaboradores pelo Centro de Estudos Avançados em Economia Aplicada (CEPEA). Estes por sua vez participam das transações nas condições de comprador, vendedor ou intermediário. Ao contatar os colaboradores os valores do preço são repassados por eles, assim como as escalas de abate, os prazos de pagamento de negócios efetivos e outras informações pertinentes (CEPEA, 2019).

Com essas informações se estabelece os critérios estatísticos que são utilizados para a composição do indicador. Os critérios considerados são: periodicidade e aleatoriedade, tentativa de contatos com os colaboradores, negócios efetivos e representatividade do negócio. Vale ressaltar que negócios em andamento, pouco representativos ou fora do padrão são descartados para a composição do indicador. Após o fechamento do indicador, o Gestor Operacional confirma a finalização dos dados para a Tecnologia da Informação (Neto, 2016).

No tratamento dos dados, valores a prazo são convertidos para Valor Presente, pela taxa de Certificado de Depósito Interbancário (CDI). O cálculo é apresentado na Figura 10.

Figura 8 - Fórmula de conversão de valores do preço a prazo para o VP da arroba de boi.
(Adaptado CEPEA, 2019)

$$valoravista = \frac{valoraprazo}{(1 + CDI_{diario})^{prazodepagamento + escaladeabate}}$$

Conforme as regiões de abate ou frigoríficos é feito uma ponderação no valor definido, com a participação de cada planta industrial no mês, que será utilizada como referência no mês seguinte. Um tratamento estatístico é feito após a definição da amostragem. Os dados são comparados com o intervalo de dois desvios padrões, e se estiverem fora do desvio, são descartados. Com todas as informações dentro de dois desvios, o Coeficiente de Variação (CV) dos dados da amostra é comparado ao valor crítico de 2,5% (a mais e a menos sobre o valor do CV). Caso o CV seja maior que o valor crítico, serão excluídos todos os negócios que correspondem ao limite inferior e superior da amostra.

Basicamente, existem três situações excepcionais que impediriam a elaboração do indicador. Um é relativo ao tamanho da amostra, e caso o total de informações coletadas no dia for menor que 20% da média móvel (considerando apenas a amostra final de cada data, dos últimos 15 dias), um colegiado interno arbitra o referido indicador. Outro é relativo ao tratamento estatístico, caso os dados resultantes após o tratamento estatístico não cumprirem os critérios de análise para a elaboração do indicador. E por último, relativo ao caso de força maior, onde o colegiado interno arbitra o referido indicador (Neto, 2016).

Diante desse contexto, para esse trabalho, os históricos de informações sobre o valor da arroba do boi, em datas retroativas, foram obtidos em páginas *web* especializadas, tais como a da Fundação de Estudos Agrários Luiz de Queiroz (FEALQ) e a BM&BOVESPA.

Etapa 2: Definir quais variáveis que compõe a precificação

Segundo Lima (2019), as variáveis utilizadas da precificação da arroba do boi podem ser agrupadas em dois fatores:

- Fatores produtivos: preço de insumos relacionado a alimentação, ao uso dos sistemas de produção, ao preço da terra, à sazonalidade, entre outros;
- Fatores de Mercado e institucionais: câmbio, embargos econômicos e de fiscalização, renda da população, preço da carne de frango, entre outros.

De acordo com a revisão da literatura, as variáveis relacionadas a alimentação dos animais, tais como, o milho e a soja têm forte influência na composição do preço da arroba do boi. Outras variáveis como o clima e a taxa de câmbio também exercem influência, podendo dizer que as variáveis até exercem influência entre elas próprias e, conseqüentemente, na composição do preço da arroba do boi.

Análises mais aprofundadas sobre as variáveis que são determinantes na definição da arroba de boi foram realizadas e coletadas. Esses dados em conjunto com o preço da arroba de boi foram utilizados na predição.

Etapa 3: Realizar as simulações com a RNA e estabelecer a arquitetura adequada.

Os testes foram realizados com Redes Neurais Artificiais do tipo *Multi Layer Perceptron*. Em função disso, algumas arquiteturas de rede foram implementadas, com neurônios na camada de entrada e na camada intermediária e, com um neurônio na camada de saída, determinando o preço de arroba de boi, na data requerida. A implementação da RNA foi implementada com três camadas. Após implementação da rede, a arquitetura foi modificada em vários testes, para obtenção de resultados, com menor erro quadrático.

Os neurônios de entrada são as variáveis relevantes na determinação da arroba de boi, sendo estas já determinadas na Etapa 2, dessa metodologia. Entretanto, com os testes da rede em simulação, verificou se algumas variáveis determinadas na etapa anterior seriam irrelevantes para a solução do problema. Em função do número de neurônios da camada de entrada (variáveis relevantes para a precificação do boi), foi definido o número de neurônios da camada intermediária e um neurônio na camada de saída, que fornece o preço da arroba de boi.

Como já mencionado, variações na implementação da rede foram realizadas quando os testes se iniciaram. Também foram testadas as variações dos parâmetros da rede, tais como, o número de padrões necessários no conjunto de treinamento, a função de ativação, o neurônio de *bias*, o algoritmo de treinamento, o modo de treinamento, a taxa de aprendizagem, a constante de momento, a inicialização dos pesos e a forma de parada de treinamento (por erro ou épocas), entre outros. Na implementação da arquitetura da rede foram utilizadas heurísticas definidas em Barbosa (2004).

Para a simulação da rede MLP foi utilizado pacote de *software* Weka (*Waikato Environment for Knowledge Analysis*), desenvolvido na Universidade de Waikato na Nova Zelândia, sendo esse de código aberto sob a licença pública *General Public License* (GNU). Esse *software* pode ser classificado como uma coleção de algoritmo de aprendizado de máquinas, que contém ferramentas para preparação, classificação, regressão e visualização de dados (Weka, 2011).

A escolha do Weka ocorreu devido à necessidade de se implementar uma Rede Neural Artificial MLP para a análise dos dados fornecidos e gerar informações estatísticas a partir dos resultados encontrados. As simulações foram realizadas e testadas para que se obtivesse um erro quadrático menor que outros implementados com RNA ou outras técnicas.

Capítulo 5

RESULTADOS

Nesse capítulo são apresentados os resultados da definição das variáveis pesquisadas e selecionadas, e a descrição de como foram realizadas as coletas de dados para serem utilizadas no modelo da RNA. Apresenta-se uma descrição sucinta das principais variáveis utilizadas (**valor_milho**, **valor_dolar_compra**, **valor_dolar_venda**, **valor_boi_australia**, **valor_real_frango**, **valor_real_bezerro**, **sazonalidade**, *IGPM*) e uma tabela com todas as variáveis pesquisadas, mas que não tiveram resultado desejado nesse trabalho.

As heurísticas utilizadas para criação dos modelos da RNA e uma visão geral do Weka também são descritas, bem como o *software* utilizado para modelagem e teste, além de todos os testes realizados.

5.1 Variáveis Pesquisadas e Coleta de Dados

Como o objetivo do trabalho é a previsão do preço da arroba do Boi Gordo, as informações referentes ao seu preço foram retiradas da base de dados do CEPEA (2021). Esses valores são os utilizados como indicadores do preço do Boi Gordo e servem de referência para os estados, como já mencionado na metodologia.

A escolha da variável **valor_milho** ocorreu devido ao fato do milho ser uma das principais fontes de nutrição do Boi Gordo, o que possivelmente impacta na composição do seu preço final. Os dados históricos do valor do milho foram extraídos da base de dados do CEPEA (2021).

O valor do dólar para compra e venda, conforme boa parte de trabalhos relacionados a este, tem uma forte influência na composição do preço da arroba do Boi Gordo. Neste trabalho teve um impacto positivo no modelo preditivo e será melhor detalhado nos testes posteriormente. O *site* do Banco Central (BACEN, 2021) possui os dados históricos diários do **valor_dolar_compra** e **valor_dolar_venda** e foram extraídos nessa plataforma para serem utilizados neste trabalho.

Até dezembro de 2019, a Austrália era o segundo país maior exportador de carne bovina no mundo, atrás apenas do Brasil (BeefPoint, 2019). E por este motivo foi considerado nos testes o valor do preço do boi da Austrália, representado pela variável **valor_boi_australia**. No *site* de Indexmundi (2020) foi realizada a coleta dos dados do preço do boi australiano, em dólar por quilograma.

O frango pode ser uma proteína substituta da carne vermelha e de acordo com o estudo dos trabalhos relacionados, foi observado que pode contribuir na formação do preço da arroba do Boi Gordo. Os históricos de valores do frango foram retirados da base de dados da CEPEA (2021) e foi simbolizado pela variável **valor_real_frango**.

Mion (2019) cita em seu trabalho a relação entre o bezerro e o Boi Gordo na avaliação dos custos da reposição do rebanho. Ou seja, quanto maior o preço do bezerro, menor é a sua quantidade adquirida ao fazer a venda de um animal pronto para o abate, podendo ocasionar em um aumento no preço do Boi Gordo. Desta forma, foi inserido a variável **valor_real_bezerro** e os dados do histórico do preço foram obtidos pelo CEPEA (2021).

A sazonalidade refere-se ao período de safra, que normalmente ocorre de janeiro até junho, com maior período de chuva e, a entressafra que começa em junho e vai até dezembro, que compreende o período mais seco. Lima (2019) em seu estudo descreve a sazonalidade como uma variável não significativa para formação do preço do Boi Gordo.

Observou-se, entretanto, que nos testes realizados nesse trabalho essa variável sazonalidade teve uma importância para aumentar a capacidade de predição do modelo. Inicialmente foram utilizadas variáveis referente às chuvas das principais praças que compõe o preço do Boi Gordo. Essas variáveis não tiveram o resultado esperado, sendo substituídas pela variável **sazonalidade**. A informação dessa variável foi usada como do tipo binária, 0 para o período de safra e 1 para o período de entressafra.

O indicador de medição da inflação IGP-M (Índice Geral de Preços do Mercado), criado pela FGV (Fundação Getúlio Vargas), é usado para medir os preços desde matérias primas até produtos relacionados ao consumidor final. Esse índice é importante para estimar o poder de compra do consumidor, e para estimar como a inflação pode interferir no preço da arroba do

boi. Os dados do Índice Geral de Preços – Mercado (IGPM) foram retirados do banco de dados da FGV.

No decorrer deste trabalho a coleta de dados e os primeiros testes foram feitos de forma gradual, conforme variáveis usadas nos trabalhos relacionados e conforme eram disponibilizados, e se eram gratuitos. Na Tabela 3 são apresentadas todas as variáveis que foram consideradas no trabalho.

Tabela 3 - Variáveis Pesquisadas para serem consideradas como entrada da RNA e a fonte de onde podem ser obtidas.

#	Variável	Frequência e obtenção dos dados	Fonte
1	data_preço_boi	diário	CEPEA
2	valor_real_milho	diário	CEPEA
3	valor_dolar_compra	diário	BACEN
4	valor_dolar_venda	diário	BACEN
5	valor_boi_australia	mensal	Indexmundi (2021)
6	valor_boi_eua	mensal	USDA
7	valor_real_frango	diário	CEPEA
8	valor_real_bezerro	diário	CEPEA
9	sazonalidade	0 - período de safra (janeiro a maio) 1 - período de entressafra (junho a dezembro)	Lima (2019)
10	IGPM	mensal	FGV
11	chuva_barretos	diário	INMET
12	chuva_bauru	diário	INMET
13	chuva_campo_grande	diário	INMET
14	chuva_presidente_prudente	diário	INMET
15	valor_dolar_milho	diário	CEPEA
16	valor_dolar_frango	diário	CEPEA
17	valor_dolar_bezerro	diário	CEPEA
18	escala_abate	trimestral	IBGE
19	valor_kq_exportacao_china	trimestral	COMEX STAT
20	doença_animal	0 - Não 1 - Sim	GOV
21	doença_humana	0 – Não 2 – Sim	GOV
22	valor_real_boi	diário	CEPEA

5.2 Heurísticas Utilizadas na Implementação da Rede MLP

No desenvolvimento desta pesquisa foram construídas várias redes, utilizando heurísticas diferentes para a definição da camada intermediária na modelagem da rede MLP. Os parâmetros e as características utilizadas nas diferentes heurísticas são descritos, como segue:

- número de Camadas: 3 considerando que caso a função tenha um número limitado de pontos ou, quando é utilizada uma função contínua, uma rede de três camadas é capaz de aprender esta função (BRAGA *et al.*, 2014);
- número da camada de Entrada: o número da camada de entrada corresponde ao número de variáveis de entrada do modelo a ser desenvolvido. Para este trabalho foram utilizados nos testes até 19 variáveis, definidas na Tabela 3 e posteriormente apresentado nos testes;
- número da camada de saída: a técnica adotada para este trabalho é a regressão que é uma forma de transformar as variáveis de entrada em uma saída numérica. Logo, na camada de saída tem um neurônio representado pela variável **valor_real_boi**;
- número de padrões necessário para treinamento: os testes foram realizados com 70% para treinamento e 30% para validação e teste do modelo de rede;
- número de neurônios na camada intermediária: para definição da camada intermediária foram testados duas heurísticas:
 - A primeira heurística (Barbosa, 2004): considera que o número de neurônios da camada intermediária depende do número de exemplos utilizado no treinamento. Dessa forma, para definir a quantidade de ligações sinápticas entre dois neurônios, encontra-se o Número Total de Pesos (NTP) que, caso possua milhares de exemplo no treinamento será definido pela Equação 1, como é o caso desse trabalho.

Equação 1

$$NTP = \frac{\text{Num de exemplos Treinamento}}{10}$$

onde:

NTP: Número Total de Pesos

Num de exemplos Treinamento: Número de exemplos de treinamento

- A segunda heurística: baseada em Hagan *et. al.* (2014) sugere uma fórmula (Equação 2) para solução de problemas de aprendizagem supervisionado, evitando o sobreajuste¹ que é uma condição que acontece quando a rede se ajusta extremamente bem aos dados de treinamento, mas por possuir uma flexibilidade excessiva, termina por prover generalizações pouco coerentes com o modelo subjacente aos dados.

Equação 2

$$Nh = \frac{Ns}{(\alpha * (Ni + No))}$$

Onde:

- Nh = Número de neurônios da camada oculta;
 - Ns = Número de exemplos do conjunto de treinamento
 - α = Valor arbitrário entre 2 e 10
 - Ni = Número de neurônios de entrada
 - No = Número de neurônios de saída
-
- algoritmo de treinamento: *backpropagation*;
 - taxa de aprendizagem: foram utilizados os valores 0.001, 0.0005 e 0.0146. Essas alterações do valor do parâmetro em alguns testes apresentaram variações nos resultados;
 - *Momentum*: Foram utilizados os valores de 0.9 e 0.99. Alguns testes apresentam uma variação nos resultados com a variação do parâmetro;
 - número de épocas: Foram utilizados valores de 5000 a 500000 nos testes desenvolvidos.

¹ É um termo usado em estatística para descrever quando um modelo estatístico é muito adequado para um conjunto de dados observado anteriormente, mas não é eficaz na previsão de novos resultados.

5.3 Implementação das RNAs

Os testes realizados têm como característica principal uma rede com 3 camadas, contudo o número de neurônios da camada de entrada e camada intermediária variam conforme o teste realizado. A camada de saída contém um neurônio em todos os testes.

O modelo foi desenvolvido para prever o preço do valor da arroba do boi gordo em um período de 90 dias futuro. Os testes na maioria das vezes foram feitos de maneira gradativa, ou seja, no modelo foram sendo adicionadas novas variáveis na camada de entrada.

Por existir uma grande quantidade de variáveis, vários testes foram realizados para chegar ao modelo de rede que mais se aproximasse ao valor de saída com menor erro (ou mais próximo do real).

Observa-se que devido ao fato deste trabalho ter sido desenvolvimento durante o período da pandemia global da COVID-19 (infecção respiratória aguda causada pelo coronavírus SARS-CoV-2, potencialmente grave, de elevada transmissibilidade e de distribuição global. 2021, Ministério da Saúde) os testes foram divididos entre o período pré-pandemia (antes de dezembro de 2019) e o período pós-pandemia (após dezembro de 2019). Nas seções posteriores serão mostrados os principais resultados das redes implementadas.

5.4 Teste Pré-Pandemia

Os testes iniciais foram para compreensão da utilização do *software* Weka, entender como funciona o simulador da rede MLP e identificar quais variáveis possuem maior relevância para predição do problema proposto. Tais testes foram realizados com as configurações padrão do Weka e, posteriormente, foram aplicadas as heurísticas destacadas na metodologia. E como já citado anteriormente os dados das variáveis para realização dos testes na pré-pandemia compreendem o período de janeiro de 2010 a dezembro de 2019, podendo haver alguma alteração no período devido a indisponibilidade dos dados.

As primeiras variáveis obtidas para a realização dos testes foram o valor do milho e o valor do dólar. A escolha em começar por essas variáveis se deve ao fato dos trabalhos relacionados as citarem como as que mais influenciam no preço da arroba do boi gordo.

No primeiro conjunto de simulação foram feitos testes com 2.488 instancias sendo 66% para o treinamento da rede e 34% para teste e validação. A Tabela 4 apresenta os resultados das simulações iniciais e a rede contendo a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 5;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda;
- período dos dados das variáveis: 04/01/2010 a 30/12/2019;
- camada intermediária: Número de neurônio de entrada dividido por dois (configuração padrão do Weka);
- valor de *Momentum*: 0.2;
- taxa de Aprendizagem: 0.3;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.642 para treinamento e 846 para teste).

Ainda na Tabela 4, são apresentados os resultados para diferentes épocas do coeficiente de correlação, erro absoluto médio, erro quadrático médio, erro relativo e erro relativo quadrático.

Tabela 4 - Resultado dos primeiros testes iniciais

Nº Épocas \ Resultados	500	1000	25.000	50.000
Coeficiente de correlação	0.9557	0.9562	0.9563	0.9563
Erro absoluto médio	4.8723	4.8473	5.0126	5.0126
Erro quadrático médio	7.9254	7.9661	8.0565	5.0126
Erro absoluto relativo	20.5511 %	20.4457 %	21.1425 %	21.1428 %
Erro quadrático relativo	29.5185 %	29.6699 %	30.0068 %	30.0067 %

Observa-se que o erro relativo dos testes realizados ficou em torno de 20%, sem mudanças significativas com o aumento do número de épocas. A taxa média de 80% pode ser considerada aceitável para alguns problemas, entretanto, para o problema proposto representa um erro elevado na definição do preço da arroba do boi para o produtor.

Na segunda etapa dos testes foi adicionado a variável `valor_boi_australia`. Esses dados representam o valor em dólar por quilograma da carne bovina na Austrália. A configuração da rede foi similar ao teste anterior, tendo as seguintes características:

- quantidade de variáveis de entrada (neurônios): 6;
- variáveis: `data_preco_boi`; `valor_real_milho`; `valor_dolar_milho`; `valor_dolar_compra`; `valor_dolar_venda`; `valor_boi_australia`;
- período dos dados das variáveis: 04/01/2010 a 30/12/2019;
- camada intermediária: número de neurônio de entrada dividido por dois;
- valor de *Momentum*: 0.2;
- taxa de Aprendizagem: 0.3;
- normalização dos atributos: verdadeiro
- método de execução do teste: *Percentagem split* 66% (1.642 para treinamento e 846 para teste);

A Tabela 5 apresentam os mesmos tipos de erros e quantidade de épocas do teste anterior.

Tabela 5 - Resultado adicionando a variável valor_boi_australia

Resultados \ N° Épocas	500	1000	25.000	50.000
Coeficiente de correlação	0.9725	0.9756	0.9756	0.9756
Erro médio absoluto	6.5433	6.5629	6.2164	6.2166
Raiz quadrada do erro médio	8.1638	7.9858	7.641	7.6412
Erro absoluto relativo	27.5992 %	27.6819 %	26.2203 %	26.2211 %
Erro quadrático relativo	30.4063 %	29.7432 %	28.4591 %	28.4598 %

A Tabela 5 mostra que a taxa de erro absoluto médio reduziu em torno de 6% com a adição da variável preço da carne bovina australiana. Como os Estados Unidos também são um grande exportador mundial de carne, a variação do preço da carne americana foi considerada e foi adicionada a variável valor_boi_eua. A configuração da rede foi:

- quantidade de variáveis de entrada (neurônios): 7;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua;
- período dos dados das variáveis: 03/01/2011 a 30/12/2019;
- camada intermediária: Número de neurônio de entrada dividido por dois;
- valor de *Momentum*: 0.2;
- taxa de aprendizagem: 0.3;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.478 para treinamento e 762 para teste);

Tabela 6 - Resultado adicionando a variável valor_boi_eua ao restante das entradas.

Resultados \ N° Épocas	500	1000	25.000	50.000
Coeficiente de correlação	0.9887	0.9888	0.9894	0.9894
Erro médio absoluto	3.2114	3.1984	2.8969	2.8921
Raiz quadrada do erro médio	4.1054	4.0995	3.8127	3.809
Erro absoluto relativo	14.5527 %	14.4939 %	13.1274 %	13.1059 %
Erro quadrático relativo	16.1007 %	16.0775 %	14.9527 %	14.9381 %

A Tabela 6 apresenta uma diminuição percentual nos erros ao adicionar a variável referente ao preço do boi nos Estados Unidos. O período da amostragem dos dados foi alterado nesse exemplo para o início de 2011 devido a disponibilidade dessa informação, mas pelo que pôde ser observado não comprometeu o resultado, por ainda restar uma grande quantidade de dados.

Em meados de 2019 houve uma epidemia de peste suína africana na China em que 1 milhão de porcos foram abatidos e 150 milhões seriam abatidos. Essa epidemia alavancou o mercado das proteínas e consequentemente a exportação da carne bovina, fazendo com o que preço da arroba do boi tivesse um aumento significativo (Revista Globo Rural, 2019).

Essa alteração no preço da arroba do boi devido a peste suína, causou um aumento percentual de erro ao fazer a modelagem da rede, como pode ser observado na Tabela 7. Esse teste não considerou o ano de 2019 e foi feito uma comparação com o teste anterior. Os parâmetros da configuração da rede são:

- quantidade de variáveis de entrada (neurônios): 7;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua;
- período dos dados das variáveis: descrito na Tabela 7;
- camada intermediária: Número de neurônio de entrada dividido por dois
- valor de *Momentum*: 0.2;
- taxa de Aprendizagem: 0.3;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

A Tabela 7 aponta para um resultado mais eficiente com a exclusão do período de 2019 no conjunto de dados. Isso é explicável, porque a quantidade de dados para o treinamento nesse período de 2019 é relativamente pequena em relação aos demais dados, não sendo suficiente para o adequado aprendizado. Foram feitos mais testes com o período até 2018 afim de melhorar o modelo de predição para a inclusão dos períodos posteriores (2019 e 2020) no conjunto de dados.

Tabela 7 - Resultado comparativo com datas diferentes

Resultados \ N° Épocas	25.000	25.000
Período da amostra de dados	03/01/2011 a 30/12/2019	03/01/2011 a 30/12/2018
Coeficiente de correlação	0.9894	0.9937
Erro médio absoluto	2.8969	0.9937
Raiz quadrada do erro médio	3.8127	2.6974
Erro absoluto relativo	13.1274 %	10.0173 %
Erro quadrático relativo	14.9527 %	11.9205 %

5.4.1 Teste utilizando a Heurística A

Depois dos testes iniciais, uma nova rodada de testes foi executada utilizando a heurística “A” definida na metodologia, e em alguns casos com variação no número de neurônios da camada intermediária e da taxa de aprendizagem. Para o próximo teste adicionou-se variáveis de chuva das principais praças de produção do gado de corte. As variáveis em que se referem o índice pluviométrico das cidades de Barretos, Bauru e Presidente Prudente foram consideradas para o experimento por fazer parte da criação do índice do preço do boi gordo do estado de São Paulo (CEPEA/ESALQ) e que exercem influência sobre os preços do restante do Brasil.

A configuração da rede teve os seguintes parâmetros:

- quantidade de variáveis de entrada (neurônios): 11;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua;

chuva_barretos; chuva_bauru; chuva_campo_grande;
chuva_presidente_prudente;

- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária: 12;
- valor de *Momentum*: 0.9;
- taxa de Aprendizagem: 0.02;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Com essa configuração houve uma melhora percentual não significativa no desempenho da rede, portanto aquém do resultado esperado. A Tabela 8 apresenta o resultado dos testes das referentes às chuvas de algumas regiões.

Tabela 8 - Resultado comparativo adicionando as variáveis referente às chuvas das principais regiões do estado de São Paulo e Campo Grande.

Nº Épocas	500	50.000	100.000
Resultados			
Coefficiente de correlação	0.9934	0.9963	0.9964
Erro médio absoluto	2.2301	1.9802	1.9882
Raiz quadrada do erro médio	2.8214	2.4328	2.4393
Erro absoluto relativo	10.6691 %	9.4734 %	9.5115 %
Erro quadrático relativo	12.4684 %	10.7512 %	10.7799 %

Os dados históricos das chuvas foram substituídos pela variável de sazonalidade, que é representada por forma binária os períodos de chuva ou seca. Essa escolha também se deve a dificuldade em trabalhar e obter os dados pluviométricos de diferentes estados brasileiros. Novos testes foram realizados e o modelo de rede novamente apresentou uma melhora com relação ao teste anterior conforme exposto na Tabela 9. A configuração da rede foi:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho;
valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua;
sazonalidade;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;

- camada intermediária: 16;
- valor de *Momentum*: 0.9;
- taxa de Aprendizagem: 0.0125;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 9 - Resultados da RNA considerando também a variável sazonalidade como característica de entrada.

Resultados \ N° Épocas	500	50.000	100.000
Coeficiente de correlação	0.9949	0.9975	0.9978
Erro médio absoluto	2.2201	1.4358	1.1256
Raiz quadrada do erro médio	2.7141	1.8355	1.5501
Erro absoluto relativo	10.6209 %	6.8687 %	5.3848 %
Erro quadrático relativo	11.9941 %	8.1114 %	6.8502 %

Realizou-se também um teste com uma variação no número de neurônios da camada intermediária, taxa de aprendizagem e número de épocas. Porém, os resultados dessa nova configuração não apresentou diferença significativa com relação aos testes anteriores. O resultado é apresentado na Tabela 10, com a rede contendo os seguintes parâmetros:

- quantidade de variáveis de entrada (neurônios): 8;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua; sazonalidade;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária: 22;
- valor de *Momentum*: 0.9;
- taxa de Aprendizagem: 0.009;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 10 - Resultado com a variável de sazonalidade e novos parâmetros de configuração.

Resultados \ N° Épocas	50.000	200.000
Coeficiente de correlação	0.9979	0.9982
Erro médio absoluto	1.2165	1.0575
Raiz quadrada do erro médio	1.5555	1.4205
Erro absoluto relativo	5.8198 %	5.0593 %
Erro quadrático relativo	6.874 %	6.2774%

No próximo modelo de rede testado teve a inserção das variáveis: valor do bezerro e valor do frango, em reais e dólares respectivamente. Nos testes realizados com essas variáveis houve desempenho com um percentual de erro menor, conforme apresentado na Tabela 11. A configuração da rede foi a seguinte:

- quantidade de variáveis de entrada (neurônios): 12;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; valor_boi_eua; sazonalidade; valor_real_bezerro; valor_dolar_bezerro; valor_real_frango; valor_dolar_frango;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária: 22;
- valor de *Momentum*: 0.9;
- taxa de Aprendizagem: 0.009;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 11 - Resultado do teste com a inserção das variáveis preço do bezerro e do frango.

Nº Épocas Resultados	50.000	200.000
Coeficiente de correlação	0.9983	0.9987
Erro médio absoluto	1.1206	0.9576
Raiz quadrada do erro médio	1.4188	1.2314
Erro absoluto relativo	5.3611 %	4.5814 %
Erro quadrático relativo	6.2702 %	5.4417 %

Dentre as várias combinações de configurações para a modelagem da rede neural, foi feito um teste incluindo a variável de Índice Geral de Preço de Mercado (*IGPM*) que está associada ao índice de inflação do país. A rede possui a seguinte configuração:

- quantidade de variáveis de entrada (neurônios): 9;
- variáveis: *data_preco_boi*; *valor_real_milho*; *valor_dolar_compra*; *valor_dolar_venda*; *valor_boi_australia*; *sazonalidade*; *valor_real_bezerro*; *valor_real_frango*; *IGPM*;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária: 22;
- valor de *Momentum*: 0.9;
- taxa de Aprendizagem: 0.009;
- normalização dos atributos: verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 12 - Resultado do teste da RNA com a adição da variável *IGPM*.

Nº Épocas Resultados	50.000	200.000
Coeficiente de correlação	0.9984	0.9983
Erro médio absoluto	0.9759	0.9349
Raiz quadrada do erro médio	1.3251	1.3057
Erro absoluto relativo	4.6687 %	4.4724 %
Erro quadrático relativo	5.8557 %	5.7703 %

O resultado na Tabela 12 mostra capacidade melhor de aprendizagem da RNA com relação aos modelos anteriormente apresentados. Inúmeros testes foram realizados com diferentes configurações dos parâmetros da rede seguindo a heurística “A”, podendo ter alguma variação dentro do que foi proposto na heurística. Os principais testes foram apresentados, e apesar de haver uma diminuição percentual dos erros nos resultados, ficou a dúvida se os resultados poderiam ser melhorados. Dessa forma, uma nova abordagem foi adotada e será descrita na seção seguinte.

5.4.2 Teste utilizando a Heurística B

Com o intuito de diminuir a taxa de erros do modelo da rede, mais testes foram realizados e a busca na melhoria da configuração da rede foi implementada com base na heurística “B”. Mais uma vez, vários testes foram realizados e muitos destes ficaram com o Erro Absoluto Relativo entre 4% e 5%, o que não seria algo diferente dos testes realizados com a heurística “A”. No entanto, houveram testes que a taxa de erro ficou abaixo dos 4% e estes serão descritos na sequência.

O resultado, de acordo com a Tabela 13, apresenta uma diminuição percentual da taxa de erro, em relação aos demais testes. Segue a configuração da rede:

- quantidade de variáveis de entrada (neurônios): 9;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária: 31 neurônios;
- valor de *Momentum*: 0.99;
- Taxa de Aprendizagem: 0.001;
- normalização dos atributos: verdadeiro;
- Método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 13 - Resultado teste com a Heurística “B” de configuração de rede.

Nº Épocas	
Resultados	150.000
Coeficiente de correlação	0.9988
Erro médio absoluto	0.8098
Raiz quadrada do erro médio	1.1143
Erro absoluto relativo	3.8743 %
Erro quadrático relativo	4.9242 %

Ainda sobre o teste, conforme Tabela 13, é importante salientar que a rede estabilizou com o número de épocas atribuído em 150.000. Números mais altos deste parâmetro causaram uma piora no desempenho, aumentando um pouco a taxa de erro do modelo de rede.

Em seguida outro conjunto de teste foi implementado com uma nova variável na camada de entrada, a *escala_abate*, que pode ter um impacto no preço da arroba do boi, conforme a oferta e procura dos animais.

A configuração da rede foi a seguinte:

- quantidade de variáveis de entrada (neurônios): 10;
- variáveis: *data_preco_boi*; *valor_real_milho*; *valor_dolar_compra*; *valor_dolar_venda*; *valor_boi_australia*; *sazonalidade*; *valor_real_bezerro*; *valor_real_frango*; *IGPM*; *escala_abate*;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária (neurônios): 36;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.001;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 14 - Resultado inserindo a variável escala_abate

Resultados	Nº Épocas
	150.000
Coeficiente de correlação	0.9989
Erro médio absoluto	0.7815
Raiz quadrada do erro médio	1.0781
Erro absoluto relativo	3.7387 %
Erro quadrático relativo	4.7643 %

Conforme Tabela 14 os resultados apresentam uma melhora percentual não muito significativa em relação ao teste anterior. Sendo assim, o próximo teste que será apresentado e que apresentou resultado abaixo dos 4% da taxa de erro absoluto foi o que incluiu as variáveis de valor_real_soja e valor_real_frango. Salieta-se que ambas as variáveis podem impactar no preço da arroba do boi. A soja por ser fonte de proteína para alimentação bovina, e o frango que é uma fonte energética alternativa a carne bovina e que possuem uma relação entre si no que se refere ao preço (LIMA, 2019).

Segue a configuração da rede:

- quantidade de variáveis de entrada (neurônios): 12
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*; escala_abate; valor_real_soja; valor_real_frango;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária (neurônios): 20;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.001;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

Tabela 15 - Resultado adicionando as variáveis valor_real_soja e valor_real_frango.

Resultados \ N° Épocas	100.000
Coeficiente de correlação	0.9989
Erro médio absoluto	0.7908
Raiz quadrada do erro médio	1.0666
Erro absoluto relativo	3.7833 %
Erro quadrático relativo	4.7137 %

Desta vez, conforme mostrado na Tabela 15, o teste teve uma diminuição percentual da taxa do erro médio absoluto em relação ao teste anterior. A última sequência de testes com uma variável diferente das demais testadas, teve o valor_ kg_exportacao_china incluída no modelo de rede. Conforme os veículos de informação como ABIEC (2021) ou sites como *farmnews* (2021) a China é um dos principais importadores de carne bovina do Brasil. Essa informação fez com que mais testes fossem realizados e com o conjunto de dados descrito a seguir:

- quantidade de variáveis de entrada (neurônios): 13;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*; escala_abate; valor_real_soja; valor_real_frango; valor_kq_exportacao_china;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- camada intermediária (neurônios): 20;
- valor de *Momentum*: 0.99;
- taxa de aprendizagem: 0.001;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66% (1.313 para treinamento e 677 para teste).

**Tabela 16 - Teste com a inclusão da variável
valor_kg_exportacao_china**

Resultados	Nº Épocas
	150.000
Coeficiente de correlação	0.9989
Erro médio absoluto	0.7791
Raiz quadrada do erro médio	1.0762
Erro absoluto relativo	3.7271 %
Erro quadrático relativo	4.756 %

Nota-se que o primeiro teste mostrado na Tabela 13 tem um resultado semelhante aos demais testes, este por sua vez contém um menor número de variáveis na camada de entrada na construção do modelo da RNA. Desta forma, este foi um dos modelos de rede adotado para a predição que será abordado posteriormente, em seção específica denominada Predição.

5.5 Teste Pandemia

Durante o desenvolvimento deste trabalho houve uma pandemia global de COVID-19, e este fato contribuiu por uma alteração considerável no resultado do modelo de rede desse período, em relação ao período em que não existia a pandemia. Os dados pandemia referem-se a um período de ano e meio (01/01/2020 a 01/06/2021) a mais no conjunto de dados que estavam sendo utilizados, e testes com este período mais recentes foram implementados.

A Tabela 17 mostra os resultados comparativos e a variação dos resultados conforme o período do conjunto de dados. Pode-se observar que o teste que considerou o ano de 2019 teve um aumento na taxa de erros considerando que neste mesmo ano teve a epidemia da peste suína africana na China e o início da pandemia de COVID-19 em dezembro. A partir de 2020 até maio de 2021 a taxa de erro estabilizou, entretanto existe um aumento da taxa de erro de quase 2%. A configuração da rede foi a seguinte:

- quantidade de variáveis de entrada (neurônios): 9;
- variáveis: `data_preco_boi`; `valor_real_milho`; `valor_dolar_compra`; `valor_dolar_venda`; `valor_boi_australia`; `sazonalidade`; `valor_real_bezerro`; `valor_real_frango`; *IGPM*;
- camada intermediária (neurônios): 31;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.001;
- normalização dos atributos: Verdadeiro;
- número de épocas: 150.000;
- método de execução do teste: *Percentagem split* 66%;

Tabela 17 - Tabela comparativa dos resultados da RNA entre períodos diferentes.

Período dos dados Resultados	03/01/2011 a 30/12/2018	03/01/2011 a 30/12/2019	03/01/2011 a 30/12/2020	03/01/2011 a 31/05/2021
Situação	Antes pandemia e peste suína	Antes pandemia e com peste suína	Antes da pandemia, com peste suína e pandemia (2020)	Antes da pandemia, com peste suína e pandemia (2021)
Coeficiente de correlação	0.9988	0.9976	0.9982	0.999
Erro médio absoluto	0.8098	1.2117	1.5769	1.5732
Raiz quadrada do erro médio	1.1143	1.7541	2.4011	2.2528
Erro absoluto relativo	3.8743 %	5.7244 %	5.6373 %	4.8206 %
Erro quadrático relativo	4.9242 %	7.2626 %	6.0133 %	4.4922 %

Outros testes foram realizados considerando um período menor de tempo para observar o comportamento da rede quando incluída as variáveis **doenca_animal** e **doenca_humana** na tentativa de representar quando houvesse uma epidemia animal ou uma pandemia global como está ocorrendo no ano de 2020 e 2021.

Primeiramente foi realizado um teste sem as variáveis referente a doença, com a rede tendo as seguintes configurações:

- período: 01/01/2017 a 31/12/2018
- quantidade de variáveis de entrada (neurônios): 9;

- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*;
- camada intermediária (neurônios): 5;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.022;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66%;

Tabela 18 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2018)

Nº Épocas	5.000	50.000	150.000
Resultados			
Coefficiente de correlação	0.862	0.8258	0.8491
Erro médio absoluto	2.8636	3.5797	3.4051
Raiz quadrada do erro médio	3.5692	4.4659	4.2843
Erro absoluto relativo	60.5894%	75.7426%	72.0478%
Erro quadrático relativo	55.976%	70.0381%	67.191%

A Tabela 18 apresenta um resultado com a taxa de erro muito alta. Isso se deve ao pouco número de exemplos para treinamento da rede. Entretanto, foi realizado outro teste semelhante agora adicionando as duas variáveis binárias *doenca_animal* e *doenca_humana* para verificar se existiria alguma alteração no resultado. Em ambas as variáveis foram atribuídas o valor de 0 que indica que não existe nenhum tipo de doença no período. A rede foi configurada da seguinte maneira:

- período: 01/01/2017 a 31/12/2018
- quantidade de variáveis de entrada (neurônios):11;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*; *doenca_animal*; *doenca_humana*;
- camada intermediária (neurônios): 5;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.034;
- normalização dos atributos: Verdadeiro;

- método de execução do teste: *Percentagem split* 66%;

Tabela 19 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2018) incluindo as variáveis *doença_animal* e *doença_humana*

Nº Épocas \ Resultados	5.000	50.000	150.000
Coefficiente de correlação	0.9535	0.9625	0.962
Erro médio absoluto	1.9412	1.4868	1.5055
Raiz quadrada do erro médio	2.4273	1.9429	1.9943
Erro absoluto relativo	41.0731%	31.4592%	31.8537%
Erro quadrático relativo	38.0668%	30.4704%	31.2771%

Como visto na Tabela 19 houve uma melhora no resultado, mas além da rede possuir poucos dados para treinamento, essas variáveis não influenciaram na rede pelo fato de estarem com o valor 0 para todo o período. Então foi incluído mais um ano no período da elaboração da rede. A nova configuração ficou a seguinte:

- período: 01/01/2017 a 31/12/2019
- quantidade de variáveis de entrada (neurônios):11;
- variáveis: *data_preco_boi*; *valor_real_milho*; *valor_dolar_compra*; *valor_dolar_venda*; *valor_boi_australia*; *sazonalidade*; *valor_real_bezerro*; *valor_real_frango*; *IGPM*; *doença_anima*; *doença_humana*;
- camada intermediária (neurônios): 7;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.018;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66%;

Tabela 20 - Resultado do teste com período o reduzido (01/01/2017 a 31/12/2019) incluindo as variáveis doença_animal e doença_humana

Nº Épocas \ Resultados	5.000	50.000	150.000
Coefficiente de correlação	0.987	0.9871	0.988
Erro médio absoluto	1.7561	1.695	1.688
Raiz quadrada do erro médio	2.8816	2.854	2.7723
Erro absoluto relativo	17.7641%	17.1457%	17.0747%
Erro quadrático relativo	17.0744%	16.9112%	16.4268%

Esse teste apresentou uma melhora em relação aos anteriores, mesmo ainda tendo poucos dados para treinamento e também um curto período de tempo considerando a variável doença_animal, sendo atribuída como 1 no mês de novembro de 2019 devido a peste suína africana ocorrida na China.

Os últimos testes realizados com esse período reduzido tiveram a base de dados considerando o período de 01/01/2017 a 31/05/2021. Desta forma, pode ser observado o comportamento da variável doença_humana. A variável doença_humana foi definida como sendo o valor 1 a partir do dia 30/01/2020, conforme o decreto da Organização Mundial da Saúde (OMS), declarando emergência de saúde pública internacional. O teste com o resultado mais relevante teve a seguinte configuração:

- período: 01/01/2017 a 31/05/2021
- quantidade de variáveis de entrada (neurônios):11;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; *IGPM*; doença_anima; doença_humana;
- camada intermediária (neurônios): 9;
- valor de *Momentum*: 0.99;
- taxa de Aprendizagem: 0.02;
- normalização dos atributos: Verdadeiro;
- método de execução do teste: *Percentagem split* 66%;

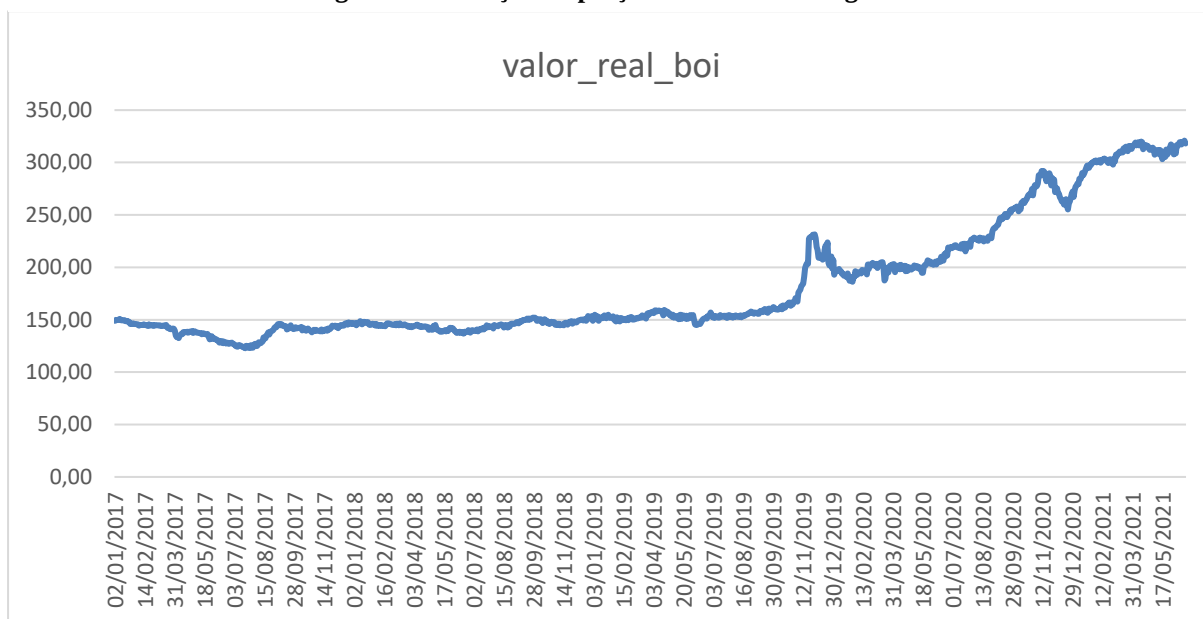
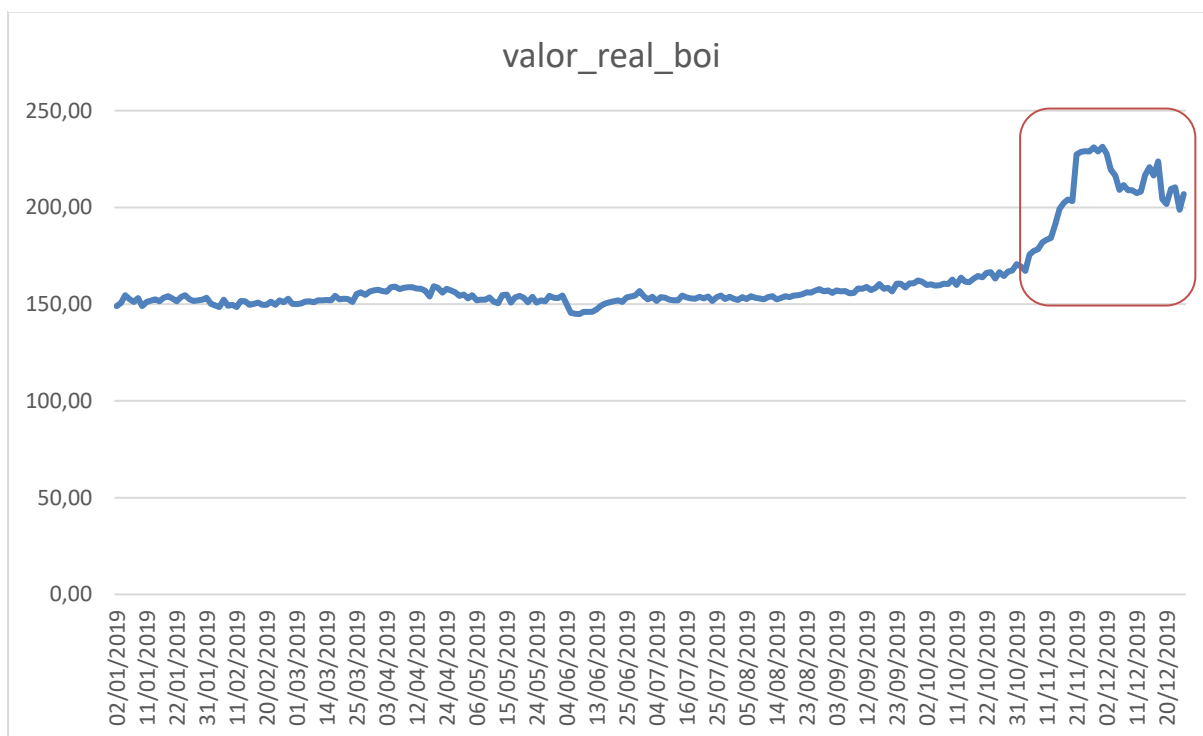
Tabela 21 - Resultado do teste com período o reduzido (01/01/2017 a 31/05/2021) incluindo as variáveis *doenca_animal* e *doenca_humana*

Resultados \ N° Épocas	5.000	50.000	150.000
Coefficiente de correlação	0.9986	0.9989	0.9988
Erro médio absoluto	2.918	2.2126	2.1343
Raiz quadrada do erro médio	3.644	2.8329	2.8076
Erro absoluto relativo	6.3259%	4.7965%	4.6269%
Erro quadrático relativo	6.5308%	5.0772%	5.0319%

A Tabela 21 mostra um resultado melhor que os anteriores considerando essa redução no *dataset* e incluído os períodos mais recentes em que o preço da arroba do boi teve um aumento significativo observado no período da pandemia do COVID-19. O conjunto de dados ainda é pequeno para a elaboração do modelo que considera uma pandemia global que acontece enquanto esse trabalho foi elaborado, mas essa informação pode ser relevante para elaboração de novos modelos no futuro.

5.6 Predição

Houve um aumento bastante significativo durante o período entre 2020 e 2021, que pode ser observado no gráfico representado na Figura 13. Entretanto, o aumento elevado do preço começou no final de 2019 conforme destacado na Figura 14. Como já citado anteriormente, o conjunto de dados para o período da pandemia ainda é pequeno para um treinamento de rede que seja satisfatório, mas pode se observar uma melhora do modelo à medida que novos conjuntos de dados eram disponíveis para esse período.

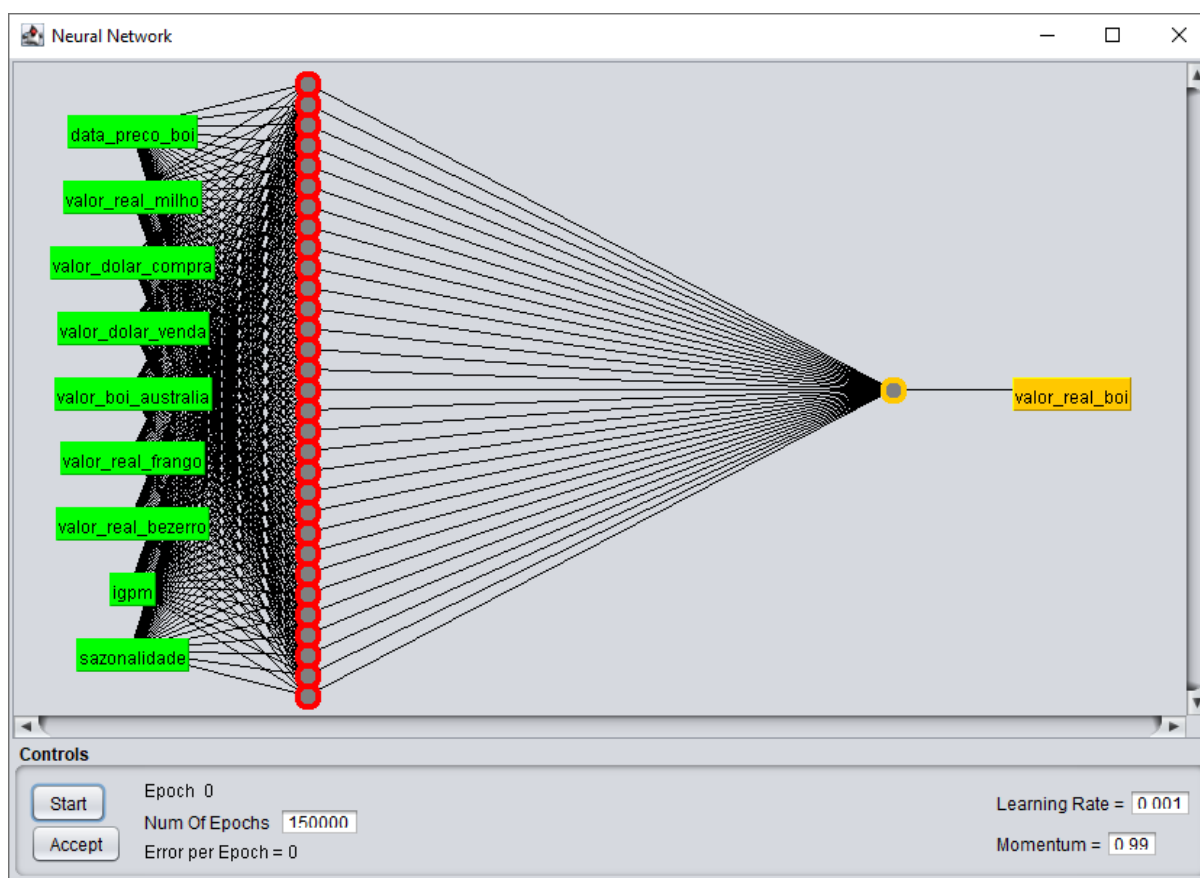
Figura 9 - Evolução do preço da arroba do boi gordo**Figura 10 - Evolução do preço da arroba do boi em 2019**

De acordo com os gráficos apresentados anteriormente, nota-se que uma das principais dificuldades para treinamento do modelo é o limite de aumento ou queda brusca do preço que pode acontecer. Também pode-se observar que a medida em que a pandemia avançava os

preços tendem a se estabilizar, porém, essa situação limítrofe entre uma alta ou queda brusca serão comentadas no capítulo sobre trabalhos futuros.

Logo após a realização de vários testes, com os principais mencionados nas seções anteriores, será apresentado o resultado da predição feita com o modelo de rede mais adequado, de acordo com a relação acurácia x quantidade de variáveis. Esta análise levou em consideração o preço da arroba do boi em até 90 dias e também períodos diferentes do *dataset*, ou seja, antes e depois do COVID-19.

Figura 11 - Imagem Rede Neural



Todas as predições a seguir possuem esta configuração no modelo de rede:

- quantidade de variáveis de entrada (neurônios): 9;
- variáveis: data_preco_boi; valor_real_milho; valor_dolar_compra; valor_dolar_venda; valor_boi_australia; sazonalidade; valor_real_bezerro; valor_real_frango; IGPM;

- camada intermediária: 31;
- valor de *Momentum*: 0.99;
- taxa de aprendizagem: 0.001;
- normalização dos atributos: Verdadeiro;
- período dos dados das variáveis: 03/01/2011 a 30/12/2018;
- número de épocas: 150.000;
- método de execução do teste: *Percentagem split* 66%;

O resultado apresentado na Figura 14; tem com data inicial para predição 02/01/19 e data final de 03/05/2019. Existe uma maior variação nos dados previstos entre o final do mês de março que coincide com a peste suína africana, ocorrida na China que desequilibrou o mercado global de carnes. A Tabela 22 contém o resultado dos 30 primeiros dias da previsão, sendo que o restante estará no Anexo A.

Figura 12 - Predição

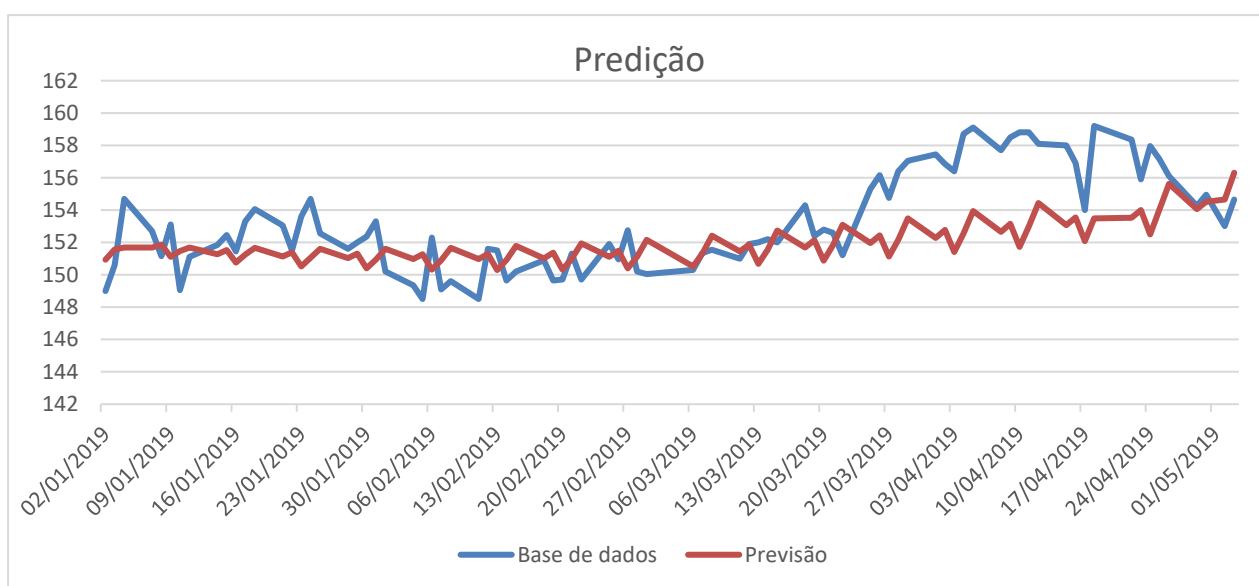


Tabela 22 - Resultado Predição dos 30 primeiros dias de predição para 90 dias futuros.

Data	Base de dados [R\$]	Previsão [R\$]	Diferença de valor absoluto do Resultado
02/01/2019	149	150,9266	-1,9266
03/01/2019	150,6	151,5684	-0,9684
04/01/2019	154,7	151,6872	3,0128
07/01/2019	152,7	151,677	1,023
08/01/2019	151,15	151,8713	-0,7213
09/01/2019	153,1	151,1001	1,9999
10/01/2019	149,05	151,4664	-2,4164
11/01/2019	151,1	151,6745	-0,5745
14/01/2019	151,85	151,2745	0,5755
15/01/2019	152,45	151,5211	0,9289
16/01/2019	151,45	150,7577	0,6923
17/01/2019	153,3	151,2447	2,0553
18/01/2019	154,05	151,6726	2,3774
21/01/2019	153,05	151,1334	1,9166
22/01/2019	151,5	151,386	0,114
23/01/2019	153,6	150,5145	3,0855
24/01/2019	154,7	151,0442	3,6558
25/01/2019	152,55	151,6115	0,9385
28/01/2019	151,6	151,0298	0,5702
29/01/2019	152	151,3105	0,6895
30/01/2019	152,35	150,39	1,96
31/01/2019	153,3	150,938	2,362

Os pesos da rede neural MLP para a predição da arroba do boi estão especificados no Anexo A.

Capítulo 6

CONSIDERAÇÕES FINAIS

6.1 Conclusão

É um desejo muito comum do ser humano querer prever o futuro. Do clima aos resultados de jogos esportivos, fica-se exposto a alguma forma de estimativa na tentativa de acertar o futuro. Muitas técnicas de predição foram desenvolvidas ao longo do tempo, e a tecnologia e modelos estatísticos tem se tornado um facilitador para esse tipo de estudo. Contudo a previsão do futuro ainda é complexa e incerta.

Este trabalho teve como objetivo a previsão do preço da arroba do boi gordo utilizando Redes Neurais Artificiais do tipo *MultiLayer Perceptron* com o auxílio de ferramenta de simulação, o *software* Weka. No decorrer da pesquisa, foi construída uma base de dados com as características consideradas importantes para formação do preço da arroba do boi, e a partir dessa base de dados e depois de vários testes houve uma seleção das variáveis que tinham um maior impacto na capacidade de predição, como visto no capítulo anterior.

O desempenho do modelo quando considerado para treinamento o período do início da pandemia de COVID-19 apresentaram um resultado não satisfatório (erro percentual de 4%). Como este trabalho foi realizado durante a pandemia e à medida que mais dados eram disponibilizados para compor o *dataset* durante a pandemia, os resultados melhoraram em torno de 0,9% em relação ao valor inicial, pois foram incluídos valores entre dezembro de 2019 e maio de 2021.

Quando não considerado o período da pandemia ou da epidemia da peste suína africana ocorrido na China no início de 2019, o modelo de rede apresentou um resultado satisfatório, com uma taxa de erro de aproximadamente **3,8%** e na média cerca de R\$ 1,00 na arroba. Como se trata de uma *commoditie* que possui uma alta volatilidade no preço, a taxa de erro foi considerada baixa.

Sobre as heurísticas utilizadas na busca do modelo de rede mais adequado, ambas se mostraram adequadas para o melhoramento do modelo proposto. A heurística B teve um desempenho um tanto quanto melhor, cerca de **1%** melhor na taxa de acerto. Porém, no decorrer dos estudos foi observado que as heurísticas servem como um norteador dos testes, e que mudanças nas configurações podem ajustar o modelo para um modelo de predição de valores.

Uma das dificuldades encontradas para elaboração desse trabalho foi a obtenção dos dados de entrada, como por exemplo a escala de abate, no qual esses dados eram disponíveis trimestralmente e que dificultou o aumento da acurácia do modelo devido ao longo período de frequência do dado. Ainda sobre a obtenção de dados, nem os bancos de dados possuem uma API (traduzido do inglês como Interface de Programação de Aplicações) que facilitaria a automatização da criação do *dataset*. Dessa forma, muitos dados foram obtidos de forma manual, através de planilhas ou em forma de texto simples.

De acordo com a análise das variáveis, foi observado que o valor do dólar teve maior influência sobre os resultados. Outras que podem ser consideradas bastante importante são referentes aos valores do frango e do milho. Faz sentido essas variáveis exercerem mais influência no preço final, visto que o frango é uma proteína substituta da carne bovina, o milho serve como base de alimento para o gado e a variação cambial interfere no preço final também. Foram excluídas as variáveis altamente correlacionadas no modelo final, exceto as dos valores do dólar compra e venda, contudo esta não apresentou estimativa errônea.

O problema da pandemia também teve um grande impacto na elaboração desse trabalho, tanto pelos dados que ainda eram insuficientes, como a alteração brusca dos preços praticados no mercado pecuarista.

Com relação ao modelo escolhido, a Rede Neural Artificial do tipo MLP, mostrou-se bastante capaz para a execução dessa tarefa, visto que é um modelo robusto com a capacidade de se adaptar conforme os dados disponíveis para treinamento.

Sobre o trabalho desenvolvido, acredita-se que é bastante significativo e que atendeu o objetivo proposto, pois apresentou valores de erro percentuais médios menores que 4%, visto que a previsão de preço diário é trabalho complexo e que envolve muitos fatores externos. Desta forma, a continuação desse trabalho pode ajudar o produtor rural a gerenciar os custos e as

receitas da produção do gado de corte e na tomada de decisão. Como, por exemplo, na baixa dos preços manter o boi no pasto até conseguir uma melhor negociação com os frigoríficos ou vender na alta.

Ainda a muito a ser explorado nesse assunto, e espera-se que esta pesquisa incentive novas pesquisas, e também seja uma fonte de conhecimento para o tema abordado.

6.2 Trabalhos Futuros

Como pôde ser observado, a questão referente à saúde pública tem um impacto no preço final da arroba do boi, tanto quando atinge a saúde humana, quanto a doença que afetam os animais e consequentemente a economia de um modo geral. Outro ponto a ser estudado no âmbito da economia são os embargos econômicos. Isso, possivelmente, também pode ter um efeito no preço da arroba bovina.

Como trabalhos futuros também propõe-se que sejam realizados testes com esse modelo proposto para prever o preço praticado em outras praças e expandir o conjunto de dados que pode influenciar o preço referente a esses estados. Apesar desta pesquisa ser feita com base com os preços praticados no estado de São Paulo, e que esse preço interfere em outros estados (CEPEA/ESALQ), ainda assim em outros estados o preço da arroba pode ser diferente.

Outra proposta seria de executar mais testes quando houver um salto ou queda brusca do preço da arroba do boi, seja por embargo ou algum fator sanitário para encontrar uma forma de correção desse problema, quando existirem dados suficientes. Esse conjunto de dados que fica entre o aumento/queda e a estabilização do preço faz com que a rede fica desbalanceada, ocasionando uma piora no desempenho. Deve-se, portanto, verificar um período adequado dos dados disponíveis para atualização do modelo, assim como a automatização ou facilitação da construção do *dataset* futuro.

Outra possibilidade seria a construção de uma aplicação *web* ou *mobile* para a visualização dos dados das variáveis e da estimativa dos preços futuros.

REFERÊNCIAS

ABIEC. **Beef Report 2019** – São Paulo 2019. Disponível <http://abiec.com.br/publicacoes/beef-report-2019/>. Acesso em: 12 mar. 2020.

Adisa, O. M. **Application of Artificial Neural Network for Predicting Maize Production in South Africa**, 2019. Disponível em: www.mdpi.com/journal/sustainability

Aguiar, H. M. **Efeitos de Variáveis Macroeconômicas no Preço do Boi Gordo do Estado de São Paulo**. Escola de Economia de São Paulo da Fundação Getúlio Vargas, 2016

Amaral, Fernando. **Aprenda Mineração de Dados – Teoria e Prática**. 1 ed. Atlas Books, 2016.

BACEM. **Banco Central do Brasil**. Disponível em: < <https://www.bcb.gov.br/>>. Acesso em: 15 de jan. 2021.

Barbosa, A T. R. **Mecanismo de Adaptação Baseado em Redes Neurais Artificiais para Sistemas Hipermídia Adaptativos**. Tese de Doutorado — Programa de Pós-graduação em Engenharia Elétrica, UFSC, Florianópolis, 2004 SC.

BeefPoint. **Portal BeefPoint**. Disponível em: < <https://www.beefpoint.com.br/cadeia-produtiva/editorial/> >. Acesso em: 25 de mar. 2020.

Braga, A. P.; Ludemir, T. B.; Carvalho, A. C. P. L. F. **Redes Neurais Artificiais: Teoria e aplicações**. 2 ed. – Rio de Janeiro: LTC, 2000.

Cezar, I. M.; Queiroz, H. P.; Thiago, L. R. L. S.; Cassales, F. L. G.; Costa, F. P. **Sistemas de produção de gado de corte no brasil: uma descrição com ênfase no regime alimentar e no abate**. Campo Grande, MS: Embrapa Gado de Corte, 2005., p. 40, 2005

CEPEA. **Centro de Estudos Avançados em Economia Aplicada**. Disponível em: <<https://www.cepea.esalq.usp.br/br>>. Acesso em: 12 de mar. 2020

Correa, A. C. M.; Leão, I. A.; Araújo, L. T.; Soares, L. A.; Souza, W. A. **Avaliação dos Preços do Boi Gordo no Estado de Goiás: Análise da Trajetória de 2008 a 2012**. Revista em Agronegócios e Meio Ambiente. V7 P. 613 – 632, 2013.

Cruz, C. D; Nascimento M. **Inteligencia Computacional Aplicada ao Melhoramento Genético** - Viçosa, 2018

ESALQ. **Metodologia do Indicador de Preços do Boi Gordo ESALQ/BM&FBOVESPA**. 2019. Disponível em: <[http://cepea.esalq.usp.br/METODOL Ind BOI GORDO Cepea-.pdf](http://cepea.esalq.usp.br/METODOL%20Ind%20BOI%20GORDO%20Cepea-.pdf)>.

Ferreira, L.; Moura, G. L.; Borenstein, D.; Fischmann, A. A. **Utilização de Redes Neurais Artificiais como estratégia de previsão de preços no contexto de Agronegócio.** Revista de Administração e Inovação, P. 6 – 26, 2011.

Globo Rural – **Revista Globo Rural**, 2019. Disponível em: <<https://revistagloborural.globo.com/Noticias/Criacao/Suinos/noticia/2019/05/febre-suina-na-china-afeta-mercados-de-proteinas-e-soja-no-mundo.html>>. Acesso em: 30/01/2021

Governo do Brasil. **O que é a Covid-19?** Site oficial do governo Brasileiro. Disponível em: <<https://www.gov.br/saude/pt-br/coronavirus/o-que-e-o-coronavirus>>. Acesso em: 23 de mai. 2020.

Grassi, M. J. S **Predição de Parto de Gado Holandês com o Uso de Redes Neurais Artificiais.** Campo Grande – Faculdade de Computação – Universidade Federal do Mato Grosso do Sul, 2017

Haykin, Simon. **Redes Neurais – Princípios e Prática.** 2 ed. Bookman Companhia, 2001.

Hawkins, D. M. **The problem of overfitting.** *Journal of chemical information and computer sciences*, ACS Publications, v. 44, n. 1, p. 1–12, 2004.

IBGE - **Instituto Brasileiro de Geografia e Estatística.** Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9203-pesquisas-trimestrais-do-abate-de-animais.html>> Acesso em: 23 mai. 2020.

Indexmundi - **Portal Indexmundi.** Disponível em: <<https://www.indexmundi.com/commodities/?commodity=beef&months=120>>. Acesso em: 25 de mar. 2020.

Lanzetta, V. B. **Uma Comparação entre Modelos de Previsão de Preços do Boi Gordo Paulista.** Piracicaba – Escola Superior de Agricultura Luiz de Queiroz, 2018

Lima, M. M. **Fatores Determinantes do Preço da Arroba do Boi Gordo.** Jaboticabal – Gestão de Organizações Agroindustriais – Universidade Estadual Paulista, 2019.

Medeiros, A. L.; Montevechi, J. A. B.; Rezende, M. L.; Reis, R. P. **Modelagem ARIMA na Previsão do Preço da Arroba do Boi Gordo.** Itajubá. - Universidade Federal de Itajubá, 2006.

Mion, R. M. R. **Previsão De Tendência De Preços Do Boi Gordo No Estado De São Paulo Utilizando Random Forest .** São Paulo – Escola de Economia de São Paulo da Fundação Getúlio Vargas, 2019

Monteiro, W. F.; Rodrigues, M. A.; Alves, A. F.; Parré, J. L. **A não efetividade do hedge para o boi gordo.** 48º Congresso Sociedade Brasileira de Economia Administração e Sociologia Rural – Campo Grande, 2010

Neto, P. M. **Sistema de Apoio a Decisão para a fase de Terminação de Bovino de Corte.** Campo Grande – Faculdade de Computação – Universidade Federal do Mato Grosso do Sul, 2016

Pereira, K A. **Análise Econômica em Sistema de Confinamento, Formação de Preços da Arroba do Boi e Suas Variáveis De Influência.** Diamantina – Universidade Federal dos Vales do Jequitinhonha e Mucuri, 2017

Sachs, R. C. C.; Pinatti, E. **Análise do comportamento dos preços do boi gordo e do boi magro na pecuária de corte paulista, no período de 1995 a 2006.** Revista de Economia e Agronegócio / Brazilian Review of Economics and Agribusiness, v. 5, 01 2007

Sanders, D. R.; Manfredo, M. R. **USDA Livestock Price Forecasts: A Comprehensive Evaluation** , *Journal of Agricultural and Resource Economics* 28(2):3 16-334, 2003

Santos , A. M.; Seixas, J. M.; Pereira, B. B.; Medronho, R. A. **Usando Redes Neurais Artificiais e Regressão Logística na Predição da Hepatite A** . Revista Epidemiol – USP . P. 117 – 126, 2005.

Schuntzemberger A. M. S. **Análise do Comportamento dos Preços do Boi Gordo da Pecuária de Corte Paranaense: Período 1994-2009 – Dissertação da** Universidade Federal do Paraná, Curitiba –, 2010

Spörl , C.; Castro, E. G.; Luchiari, A. **Aplicação de Redes Neurais Artificiais na Construção de Modelos de Fragilidade Ambiental.** Revista do Departamento de Geografia – USP . P. 113 – 135, 2011.

Weka. **Machine Learning Group at the University of Waikato**, 2011. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>> Acesso em: 01 dez. 2019.

ANEXO A

TABELA COM OS PESOS DA REDE MLP

Tabela 23 - Resultado dos neurônios intermediários para o neurônio de saída da MLP

Função Linear Node 0	
Inputs	Weights
Threshold	0.3722869460536923
Node 1	-0.34921562998119793
Node 2	1.182060108478678
Node 3	-0.8349452383272886
Node 4	-2.559358046182132
Node 5	3.5482646446828796
Node 6	0.9043331391551689
Node 7	-0.6757844528273225
Node 8	0.445050988143554
Node 9	0.4043134529823892
Node 10	0.7032713050311704
Node 11	0.5478014625253977
Node 12	-1.2711385287466228
Node 13	-1.073882983067217
Node 14	1.002554570424627
Node 15	0.39078032045545563
Node 16	0.712390199292053
Node 17	-0.4931795493098122
Node 18	-0.4926445889742003
Node 19	-0.6517813753210396
Node 20	-1.0785285870362151
Node 21	0.5984086376093023
Node 22	-0.8828253166818376
Node 23	-0.594345631906411
Node 24	-1.0209725526365905
Node 25	-0.9131966653722029
Node 26	-0.18972597614857115
Node 27	-0.2570596979224492
Node 28	-0.421436272201646
Node 29	0.8346369317168394
Node 30	-0.417675434484279
Node 31	-1.0972540360849725

Tabela 24 - Valor dos Pesos das entradas para os neurônios intermediários da MLP

Função Sigmoidal neurônio intermediário 1	
Inputs	Weights
Threshold	-8.54303018748717
Attrib data_preco_boi	7.314409370762095
Attrib valor_real_milho	3.912138294715133
Attrib valor_dolar_compra	-2.200248566609796
Attrib valor_dolar_venda	-2.201399130980901
Attrib valor_boi_australia	0.628181029618652
Attrib valor_real_frango	7.083196576935996
Attrib valor_real_bezerro	-10.329341086667602
Attrib igpm	-3.7837839510602285
Attrib sazonalidade	4.981493484796057
Função Sigmoidal neurônio intermediário 1	
Inputs	Weights
Threshold	-9.725580715641692
Attrib data_preco_boi	4.379591101258141
Attrib valor_real_milho	-0.5064565880475277
Attrib valor_dolar_compra	-0.7644371769360122
Attrib valor_dolar_venda	-0.11473664796518328
Attrib valor_boi_australia	-5.277879099235118
Attrib valor_real_frango	4.504302179016275
Attrib valor_real_bezerro	-2.485311982206667
Attrib igpm	-7.121504584827721
Attrib sazonalidade	1.925440819379973
Sigmoid Node 3	
Inputs	Weights
Threshold	-12.363053377093086
Attrib data_preco_boi	13.13070279403455
Attrib valor_real_milho	4.262111017962421
Attrib valor_dolar_compra	0.23218794650260938
Attrib valor_dolar_venda	0.21438572731612038
Attrib valor_boi_australia	-1.7800444163688576
Attrib valor_real_frango	2.4045214827303
Attrib valor_real_bezerro	1.7971212703024262
Attrib igpm	-5.053176263881331
Attrib sazonalidade	2.043735559536045

Sigmoid Node 4	
Inputs	Weights
Threshold	-4.644811239830554
Attrib data_preco_boi	-7.275837014081531
Attrib valor_real_milho	1.1112664347026304
Attrib valor_dolar_compra	0.9851881893065404
Attrib valor_dolar_venda	0.9641880918018348
Attrib valor_boi_australia	1.2682920728225862
Attrib valor_real_frango	3.641496342568623
Attrib valor_real_bezerro	1.7688838919500394
Attrib igpm	-3.1355763007660236
Attrib sazonalidade	0.8523254646018312
Sigmoid Node 5	
Inputs	Weights
Threshold	-6.2089114959225675
Attrib data_preco_boi	3.7857597111147303
Attrib valor_real_milho	-2.1581334176288025
Attrib valor_dolar_compra	-0.6747225840395678
Attrib valor_dolar_venda	0.020193249514815954
Attrib valor_boi_australia	3.8471559263392634
Attrib valor_real_frango	0.5542898511064592
Attrib valor_real_bezerro	-2.845817905622771
Attrib igpm	0.23042181085683278
Attrib sazonalidade	-2.0228326685174305
Sigmoid Node 6	
Inputs	Weights
Threshold	-6.138039439493604
Attrib data_preco_boi	-5.781667180993543
Attrib valor_real_milho	3.501099344092697
Attrib valor_dolar_compra	-0.8431564676533838
Attrib valor_dolar_venda	-0.7836263719294795
Attrib valor_boi_australia	-1.524971667424295
Attrib valor_real_frango	1.6695074808224577
Attrib valor_real_bezerro	1.1400039498702337
Attrib igpm	1.0957399697266765
Attrib sazonalidade	-2.1539650274565427

Sigmoid Node 7	
Inputs	Weights
Threshold	-2.7853217692819463
Attrib data_preco_boi	2.7556351989627443
Attrib valor_real_milho	4.672381481121558
Attrib valor_dolar_compra	-3.1987452716699005
Attrib valor_dolar_venda	-3.1804148173342313
Attrib valor_boi_australia	-0.8406573473006638
Attrib valor_real_frango	-0.518851487630295
Attrib valor_real_bezerro	3.6406074956363716
Attrib igpm	4.031998801072579
Attrib sazonalidade	2.4931296087952575
Sigmoid Node 8	
Inputs	Weights
Threshold	11.666954198947408
Attrib data_preco_boi	5.479196788629092
Attrib valor_real_milho	12.939326120079317
Attrib valor_dolar_compra	3.144096663059628
Attrib valor_dolar_venda	3.3259301696465338
Attrib valor_boi_australia	1.8884142596378337
Attrib valor_real_frango	-3.6833299072385834
Attrib valor_real_bezerro	12.589793560886765
Attrib igpm	14.98309796198636
Attrib sazonalidade	-5.007192444235641
Sigmoid Node 9	
Inputs	Weights
Threshold	-1.5504927801778132
Attrib data_preco_boi	-0.18253462162806838
Attrib valor_real_milho	3.111881285432467
Attrib valor_dolar_compra	0.3255111058883541
Attrib valor_dolar_venda	0.4335371545839571
Attrib valor_boi_australia	5.557687614498382
Attrib valor_real_frango	2.563991822576053
Attrib valor_real_bezerro	4.054291383993311
Attrib igpm	-1.0563242967696023
Attrib sazonalidade	-1.5667080962113147

Tabela 25 - Valor dos Pesos da MLP

Sigmoid Node 10		Sigmoid Node 13		Sigmoid Node 16
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -5.732029377727149		Threshold -3.154704518829767		Threshold -3.9592013548339375
Attrib data_preco_boi -5.175364121845201		Attrib data_preco_boi 0.17014619627342065		Attrib data_preco_boi 3.8600502504069047
Attrib valor_real_milho -3.3306828309830974		Attrib valor_real_milho 1.2012963036021804		Attrib valor_real_milho 0.9902881195613039
Attrib valor_dolar_compra -1.1512710146650476		Attrib valor_dolar_compra 0.2326316584146013		Attrib valor_dolar_compra 1.3651131527466032
Attrib valor_dolar_venda -1.242643087799116		Attrib valor_dolar_venda 0.18966093278758983		Attrib valor_dolar_venda 1.3705022832945575
Attrib valor_boi_australia 4.550186587562389		Attrib valor_boi_australia 2.212552727457924		Attrib valor_boi_australia 2.6723907583262823
Attrib valor_real_frango 1.8757788335905652		Attrib valor_real_frango -0.5631151074453828		Attrib valor_real_frango 1.1583056434921155
Attrib valor_real_bezerro -0.24087911081207067		Attrib valor_real_bezerro -0.39603128836842255		Attrib valor_real_bezerro 0.08757526677943818
Attrib igpm 1.6625165190935252		Attrib igpm 0.8301651290046169		Attrib igpm 0.920128398780412
Attrib sazonalidade 0.37346018295160166		Attrib sazonalidade 1.216498778079138		Attrib sazonalidade 0.9535176913276096
Sigmoid Node 11		Sigmoid Node 14		Sigmoid Node 17
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -8.045298343086179		Threshold -9.571256323115232		Threshold -0.19732577529386996
Attrib data_preco_boi 9.200722984320706		Attrib data_preco_boi -0.47395171248261797		Attrib data_preco_boi -0.7667515160739327
Attrib valor_real_milho 6.734644984321738		Attrib valor_real_milho 1.091746605588748		Attrib valor_real_milho -0.7418381364781558
Attrib valor_dolar_compra -0.6285049261076491		Attrib valor_dolar_compra -3.022668148518686		Attrib valor_dolar_compra 1.164463982513664
Attrib valor_dolar_venda -0.6871126429859481		Attrib valor_dolar_venda -3.066182007476718		Attrib valor_dolar_venda 1.0262065330871561
Attrib valor_boi_australia -6.5994421749948		Attrib valor_boi_australia -6.6363651744279775		Attrib valor_boi_australia 4.442709374225881
Attrib valor_real_frango -3.7935424960144517		Attrib valor_real_frango -6.892404336572945		Attrib valor_real_frango 1.2697702524867078
Attrib valor_real_bezerro 2.4688960972851266		Attrib valor_real_bezerro 7.621623068439741		Attrib valor_real_bezerro -0.8532551056074104
Attrib igpm 6.363248913427683		Attrib igpm 0.030360153799473995		Attrib igpm 1.76380212000065
Attrib sazonalidade -2.408476960104248		Attrib sazonalidade 2.174848953296819		Attrib sazonalidade -0.1966888987337284
Sigmoid Node 12		Sigmoid Node 15		Sigmoid Node 18
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -9.316354907679106		Threshold 0.41318617562233595		Threshold -1.5534927305431878
Attrib data_preco_boi 10.17879136064926		Attrib data_preco_boi -1.8607484130700993		Attrib data_preco_boi -3.177536792885918
Attrib valor_real_milho -2.822883449367976		Attrib valor_real_milho 7.316309275159714		Attrib valor_real_milho 5.890841457241539
Attrib valor_dolar_compra 1.6390778749275012		Attrib valor_dolar_compra -0.9738915635222544		Attrib valor_dolar_compra 0.8724562250171916
Attrib valor_dolar_venda 1.1403927815932433		Attrib valor_dolar_venda -1.0158885525845234		Attrib valor_dolar_venda 0.8323635009545168
Attrib valor_boi_australia 9.141071661656001		Attrib valor_boi_australia 3.707998546436235		Attrib valor_boi_australia 5.013002034274005
Attrib valor_real_frango 1.9242554961477254		Attrib valor_real_frango -4.994942174104256		Attrib valor_real_frango -1.26084248050671
Attrib valor_real_bezerro -3.3318081134440654		Attrib valor_real_bezerro 0.6583015504275369		Attrib valor_real_bezerro -1.7786118668550752
Attrib igpm 1.8105279260742133		Attrib igpm -5.032765333802835		Attrib igpm 4.963349458664027
Attrib sazonalidade -2.7495322496173817		Attrib sazonalidade 1.9105423747027828		Attrib sazonalidade 0.598227090786828

Tabela 26 - Valor dos Pesos da MLP

Sigmoid Node 19		Sigmoid Node 22		Sigmoid Node 25
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -6.0655358295832364		Threshold -8.536728547719136		Threshold -3.4051813978910648
Attrib data_preco_boi 3.249842452064371		Attrib data_preco_boi -7.395124070352722		Attrib data_preco_boi 0.4463699394166376
Attrib valor_real_milho 4.825792243943786		Attrib valor_real_milho -2.847320815376414		Attrib valor_real_milho 0.9366723169715211
Attrib valor_dolar_compra 0.3187898911449415		Attrib valor_dolar_compra 0.6891716078254719		Attrib valor_dolar_compra -0.6656043473236366
Attrib valor_dolar_venda 0.10409688850817085		Attrib valor_dolar_venda 0.6951173139152201		Attrib valor_dolar_venda -0.7115570098446624
Attrib valor_boi_australia -4.72210351796807		Attrib valor_boi_australia 4.226657319249278		Attrib valor_boi_australia -1.3352058000706903
Attrib valor_real_frango -4.042723261360552		Attrib valor_real_frango 0.6805882568004206		Attrib valor_real_frango 1.129523622429353
Attrib valor_real_bezerro 3.2360663772219067		Attrib valor_real_bezerro -1.6416282676163505		Attrib valor_real_bezerro -0.18200502434456134
Attrib igpm -1.569839876444835		Attrib igpm 7.396956127823778		Attrib igpm -0.34850111407044854
Attrib sazonalidade -2.565462088361496		Attrib sazonalidade -0.7827206219445222		Attrib sazonalidade -0.7901903735408566
Sigmoid Node 20		Sigmoid Node 23		Sigmoid Node 26
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -2.513151653939832		Threshold -10.726027514445692		Threshold -7.1017837216950666
Attrib data_preco_boi -12.59908515279231		Attrib data_preco_boi 3.9546600659637727		Attrib data_preco_boi 4.2687256974851575
Attrib valor_real_milho 2.16951859861134		Attrib valor_real_milho -6.966968501695715		Attrib valor_real_milho -0.4242812691998696
Attrib valor_dolar_compra -0.42934324786735734		Attrib valor_dolar_compra 2.407091025706211		Attrib valor_dolar_compra 0.8494250546199392
Attrib valor_dolar_venda -0.39161775487244094		Attrib valor_dolar_venda 2.4215350224038144		Attrib valor_dolar_venda 0.8857704917264768
Attrib valor_boi_australia -1.514210000387437		Attrib valor_boi_australia -4.442574883917265		Attrib valor_boi_australia -2.2990339598668723
Attrib valor_real_frango -1.9164435227071606		Attrib valor_real_frango 5.494924431158969		Attrib valor_real_frango -10.225321136876001
Attrib valor_real_bezerro 0.5049718792638301		Attrib valor_real_bezerro -4.577162678904645		Attrib valor_real_bezerro 0.38151812747655567
Attrib igpm 0.578235879283491		Attrib igpm -6.974872901720466		Attrib igpm 1.2355573553311603
Attrib sazonalidade -0.16157679364957522		Attrib sazonalidade 0.008592929866795633		Attrib sazonalidade 1.7454978432127086
Sigmoid Node 21		Sigmoid Node 24		Sigmoid Node 27
Inputs Weights		Inputs Weights		Inputs Weights
Threshold -6.670421141318378		Threshold -9.169957698248087		Threshold -8.647163976824174
Attrib data_preco_boi -0.041723706421756546		Attrib data_preco_boi -4.296501925816977		Attrib data_preco_boi -0.5612221443075758
Attrib valor_real_milho -0.7719589377410739		Attrib valor_real_milho 5.290510865613294		Attrib valor_real_milho -1.145421558035348
Attrib valor_dolar_compra -0.41705621948040694		Attrib valor_dolar_compra -3.271849965688561		Attrib valor_dolar_compra -4.7591295105500535
Attrib valor_dolar_venda -0.4703897864444313		Attrib valor_dolar_venda -3.2630124968401697		Attrib valor_dolar_venda -4.738916279432369
Attrib valor_boi_australia 2.2775840967160876		Attrib valor_boi_australia -5.247316119089758		Attrib valor_boi_australia 5.990129135911361
Attrib valor_real_frango -2.3606905085025693		Attrib valor_real_frango -5.548415131571384		Attrib valor_real_frango -6.413141524685186
Attrib valor_real_bezerro -0.8999054606874554		Attrib valor_real_bezerro 6.513652780744378		Attrib valor_real_bezerro -0.452287519852204
Attrib igpm 9.754718041545335		Attrib igpm 0.039423301438995556		Attrib igpm 5.071366749224239
Attrib sazonalidade -1.413124755485489		Attrib sazonalidade 1.9276461592547258		Attrib sazonalidade -1.1994616025459068

Tabela 27 - Valor dos Pesos da MLP

Sigmoid Node 28		Sigmoid Node 30
Inputs Weights		Inputs Weights
Threshold -8.096006062186314		Threshold -1.8002841465796522
Attrib data_preco_boi 5.8801776448206144		Attrib data_preco_boi -4.2315039811930495
Attrib valor_real_milho -2.7852362383172586		Attrib valor_real_milho 4.391155276260726
Attrib valor_dolar_compra -2.5482706677499936		Attrib valor_dolar_compra -0.4431023387603744
Attrib valor_dolar_venda -2.5133880296998545		Attrib valor_dolar_venda -0.5812583581151973
Attrib valor_boi_australia -0.596647875392085		Attrib valor_boi_australia 7.227371274926826
Attrib valor_real_frango 2.6139689271009043		Attrib valor_real_frango 0.48951999480095165
Attrib valor_real_bezerro 1.7927114115673983		Attrib valor_real_bezerro 1.8627655145464668
Attrib igpm 2.223946658140173		Attrib igpm -0.892336353735214
Attrib sazonalidade -2.9587715931342164		Attrib sazonalidade -3.487870601716748
Sigmoid Node 29		Sigmoid Node 31
Inputs Weights		Inputs Weights
Threshold -2.172084832603626		Threshold -6.522627267473086
Attrib data_preco_boi -0.18736997107322523		Attrib data_preco_boi -0.716061312373145
Attrib valor_real_milho 6.12735959387944		Attrib valor_real_milho -3.2450957579268396
Attrib valor_dolar_compra -2.341385749632155		Attrib valor_dolar_compra -0.27474363920945644
Attrib valor_dolar_venda -2.3894760093386656		Attrib valor_dolar_venda -0.30582131538621293
Attrib valor_boi_australia 0.5432695604233337		Attrib valor_boi_australia 0.9947649366252267
Attrib valor_real_frango 1.1548622453214468		Attrib valor_real_frango -0.25001963518817144
Attrib valor_real_bezerro 1.9891661212878713		Attrib valor_real_bezerro 0.3562575109536535
Attrib igpm 3.602415470398105		Attrib igpm -4.035532475381386
Attrib sazonalidade 2.3250606787004227		Attrib sazonalidade 1.0140563876961326