

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

GABRIELA KAROLINE MICHELON

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL NA
AGRICULTURA DE PRECISÃO PARA ESTIMAR A
PRODUTIVIDADE DA SOJA**

TRABALHO DE CONCLUSÃO DE CURSO

MEDIANEIRA

2016

GABRIELA KAROLINE MICHELON

**APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL NA
AGRICULTURA DE PRECISÃO PARA ESTIMAR A
PRODUTIVIDADE DA SOJA**

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

Orientador: Prof. Dr. Paulo Lopes de Menezes

Co-orientador: Prof. Dr. Arnaldo Candido Junior

MEDIANEIRA

2016



TERMO DE APROVAÇÃO

APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL NA AGRICULTURA DE PRECISÃO PARA ESTIMAR A PRODUTIVIDADE DA SOJA

Por

GABRIELA KAROLINE MICHELON

Este Trabalho de Conclusão de Curso foi apresentado às 10:20h do dia 18 de novembro de 2016 como requisito parcial para a obtenção do título de Bacharel no Curso de Ciência da Computação, da Universidade Tecnológica Federal do Paraná, Câmpus Medianeira. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Paulo Lopes de Menezes
UTFPR - Câmpus Medianeira

Prof. Claudio Leones Bazzi
UTFPR - Câmpus Medianeira

Prof. Jorge Aikes Junior
UTFPR - Câmpus Medianeira

Prof. Arnaldo Candido Junior
UTFPR - Câmpus Medianeira

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

RESUMO

MICHELON, Gabriela Karoline. APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL NA AGRICULTURA DE PRECISÃO PARA ESTIMAR A PRODUTIVIDADE DA SOJA. 99 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2016.

A agricultura de precisão é uma técnica de gerenciamento de áreas agrícolas que auxilia na descoberta dos fatores que causam variabilidade espacial do solo e da planta. Com as informações obtidas pela prática da agricultura de precisão, é possível saber quais pontos da lavoura necessitam de alguma correção do solo para aumentar a produtividade, o que consequentemente evita desperdícios de insumos agrícolas levando à diminuição dos gastos de produção e da degradação ambiental. As técnicas de agricultura de precisão têm sido favorecidas pelo surgimento de equipamentos eletrônicos, softwares e tecnologias da informação. Dentre as muitas tecnologias computacionais, estão as técnicas de inteligência artificial, as quais têm sido amplamente exploradas para auxiliar na produção agrícola como técnicas que permitem, através de dados característicos da planta e/ou do solo, prever a produtividade das culturas. Portanto, neste trabalho escolheu-se aplicar técnicas de inteligência artificial na cultura da soja, por ser uma cultura muito importante para consumo humano e animal, e como fonte alternativa de biocombustível no mundo todo. Propõe-se com este estudo, por meio dos macronutrientes da folha da soja e por meio de técnicas de inteligência artificial, prever a produtividade da cultura da soja, de modo que o produtor esteja em tempo de poder melhorar a produtividade da cultura plantada. As técnicas de inteligência artificial utilizadas neste estudo foram redes neurais artificiais e máquinas de vetores de suporte para regressão. Os dados da folha da soja utilizados foram Nitrogênio, Fósforo, Potássio, Cálcio e Magnésio coletados em três estádios do desenvolvimento da planta em duas áreas amostrais. Como resultado, se obteve um bom modelo de rede neural artificial capaz de explicar 74% dos dados reais, utilizando todos os nutrientes da folha da soja da primeira coleta dos dados nas duas áreas amostrais. Realizou-se treinamento das áreas amostrais separadamente, e na segunda coleta dos nutrientes da soja da área A, obteve-se um segundo melhor desempenho dentre todos os testes realizados com redes neurais artificiais e máquinas de suporte de vetores explicando 61% dos dados reais. A seleção de atributos CfsSubsetEval foi realizada para buscar um modelo de predição da produtividade da soja utilizando menos nutrientes da folha, e consequentemente, diminuir o custo de análises foliares para aplicação da técnica. No entanto, o melhor modelo obtido foi da segunda coleta dos nutrientes da folha da soja da área B, por meio de uma máquina de suporte de vetores explicando 58% dos dados reais. Portanto, este trabalho encontrou bons modelos de predição da produtividade da soja, utilizando para isto, apenas os nutrientes da folha.

Palavras-chave: regressão, redes neurais artificiais, máquinas de vetores de suporte, soja

ABSTRACT

MICHELON, Gabriela Karoline. APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES IN PRECISION AGRICULTURE TO ESTIMATE THE SOYBEAN PRODUCTIVITY. 99 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2016.

The precision agriculture is a management technique of agricultural areas that aids in the detection of the factors that causes spatial variability of soil and plant. With the information obtained by the practice of precision agriculture, it is possible to know which points of the tillage need some correction of soil to increase the productivity, what consequently avoids waste of agricultural inputs, leading to decreased costs of production and environmental degradation. The precision agriculture techniques have been favored by the appearance of electronic equipment, softwares and information technologies. Among the many computer technologies, there are the artificial intelligence techniques, which have been widely exploited to assist in agricultural production as techniques that allows, using characteristic data of the plant and/or soil, predict crop yields. Therefore, in this study was chose to apply artificial intelligence techniques in soybean, because it is a very important culture for the human and animal consumption, and as a biofuel alternative in all the world. It is proposed to this study, through the soybean leaf macronutrients and through artificial intelligence techniques to predict the soybean yield, so that the producer is in time to improve the productivity of the crop planted. The artificial intelligence techniques that were used in this study were artificial neural networks and support vector machines for regression. The soybean sheet data were Nitrogen, Phosphorus, Potassium, Calcium and Magnesium collected in three stadiums of plant development at two sites. The results are that was obtained a good predictivity model of artificial neural network, which is able to explain 74% of the real data, using the soybean leaf nutrients of the first data collection in the two sites. It conducted training of sampling areas separately, and in the second collection of nutrients from soybean area A, obtained a second best performance of all the tests with artificial neural networks and support vector machines, explaining 61% of the real data. Selecting CfsSubsetEval attributes was performed to seek a prediction model of soybean yield using fewer nutrients of leaf, and consequently reduce the cost of leaf analysis for the technical application. However, the best model is obtained from the second collection soybean leaf nutrients from area B by means of a support vector machine, explaining 58% of real data. Therefore, this study found good models of soybean yield prediction, using to this, only the leaf nutrients.

Keywords: regression, artificial neural networks, support vector machines, soybean

Dedico meu trabalho aos meus amáveis pais Lenir e Claudio, às minhas queridas irmãs Catia, Carina e Fernanda pelo amor, incentivo e apoio incondicional.

A Deus, por ter me dado forças, saúde e perseverança para concluir o curso de Ciência da Computação e não desistir jamais dos meus objetivos.

AGRADECIMENTOS

A Deus, por me ajudar a superar os obstáculos e não desanimar diante das derrotas.

A minha família, pela paciência e apoio em todos os momentos.

Ao meu orientador de Iniciação Científica Claudio Leones Bazzi, pelos anos de pesquisa juntos e pelos seus enormes conhecimentos passados a mim.

Ao meu orientador do Trabalho de Conclusão do Curso Paulo Lopes de Menezes, por ter me auxiliado na realização do trabalho e ter fornecido informações indispensáveis para o desenvolvimento deste.

Ao meu coorientador do Trabalho de Conclusão do Curso Arnaldo Cândido Junior, pelas suas contribuições e pela sua atenciosidade nas revisões do trabalho.

Ao professor Jorge Aikes Junior, por ter ensinado boas práticas de realização do Trabalho de Conclusão do Curso.

Aos demais professores, por serem íntegros e darem o melhor de si para ensinar aos seus alunos, mostrando-se sempre dispostos a tirar dúvidas e a disponibilizarem seu tempo.

Ao proprietário das áreas experimentais por permitir a realização de trabalhos utilizando os dados agrícolas. Também a todos os colegas da equipe que auxiliaram e auxiliam nas coletas dos dados amostrais, pois sem a colaboração de cada um não seria possível realizar os trabalhos aplicados na admirável área agrícola.

Aos meus colegas, em especial Felipe, Marcela, Valéria e Carlos, que ao longo destes quatro anos fizeram parte desta trajetória, cada qual com seu carisma e companheirismo, tornando meus dias mais alegres.

Aos meus amigos, por serem compreensivos e incentivadores do meu esforço, mesmo muitas vezes não podendo compartilhar muito do meu tempo e lhes dar a atenção merecida.

Ao meu companheiro de todas as horas, Jean Carlos Grandó, por sempre me apoiar a conquistar meus objetivos e por não me deixar desanimar diante dos obstáculos.

Enfim, meu eterno MUITO OBRIGADA a todos que contribuíram de alguma forma para a realização deste sonho.

”Nos dias de hoje, cada vez mais, acentua-se a necessidade de ser forte. Mas não há uma fórmula mágica que nos faça chegar à força sem que antes tenhamos provado a fraqueza.”

Padre Fábio de Melo

LISTA DE SIGLAS

AM	Aprendizado de Máquina
AP	Agricultura de Precisão
API	Application Programming Interface
ART	<i>Adaptive Resonance Theory</i>
C	Carbono
Ca	Cálcio
CFS	Correlation based Feature Selection
Co	Cobalto
CTC	Capacidade de Troca de Cátions
Cu	Cobre
CV	Coefficiente de Variação
Fe	Ferro
GIS	<i>Geographic Information System</i>
GNSS	<i>Global Navigation Satellite Systems</i>
IA	Inteligência Artificial
IAPAR	Instituto Agrônomo do Paraná
JavaNNS	Java Neural Network Simulator
K	Potássio
LMS	<i>Least Mean Square</i>
Mg	Magnésio
MLP	<i>Multilayer Perceptron</i>
Mn	Manganês
MO	Matéria Orgânica
MSE	<i>Mean squared error</i>
N	Nitrogênio
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
P	Fósforo
pH	Potencial de Hidrogeniônico
RBF	<i>Radial Basis Function</i>
REIP	<i>Red Edge Inflection Point</i>
RNA	Rede Neural Artificial
S	Enxofre
SNNS	Stuttgart Neural Network Simulator
SVM	Support Vector Machine
SVR	<i>Support Vector Regression</i>
SVs	<i>Support Vectors</i>
TA	Teor de Argila
TAE	Teoria de Aprendizado Estatístico
ton ha ⁻¹	Toneladas por hectare
UTFPR	Universidade Tecnológica Federal do Paraná
V%	Saturação por bases

WEKA Waikato Environment for Knowledge Analysis
Zn Zinco

LISTA DE SÍMBOLOS

x_i	Sinais de entrada
w_{ki}	Pesos sinápticos
Σ	Combinador linear
b_k	Limiar de ativação
v_k	Potencial de ativação
$\varphi(\cdot)$	Função de ativação
y_k	Sinal de saída
β	Constante de inclinação
c	Centro da função gaussiana
σ	Desvio padrão da função gaussiana
$w_j(t)$	Peso da j-ésima conexão de entrada em um determinado tempo
t	Tempo
η	Taxa de aprendizado
x_i^j	Valor de um vetor
$f(x_i)$	Saída produzida pela rede
y_i	Saída desejada
ε_i	Variáveis de folga
$u \cdot a$	Produto escalar entre vetores
u	Valor contido no vetor de treinamento
d	Termo compensador
Z	Espaço de entradas
γ	Espaço de características
Φ	Mapeamento
K	Matriz
p	Potência
∂	Amplitude
$h(x)$	Função
ε	Desvio
ξ_i	Variáveis de folga
α	Variável de Lagrange
ϖ	Função Kernel
m	Valor normalizado
l	Valor original

LISTA DE FIGURAS

FIGURA 1	– Hierarquia de aprendizado	21
FIGURA 2	– Neurônio biológico simplificado	25
FIGURA 3	– Organização estrutural dos níveis no cérebro	26
FIGURA 4	– Modelo de neurônio artificial	27
FIGURA 5	– Função de ativação degrau	28
FIGURA 6	– Função de ativação degrau bipolar	28
FIGURA 7	– Função de ativação rampa simétrica	29
FIGURA 8	– Função de ativação logística	29
FIGURA 9	– Função de ativação tangente hiperbólica	30
FIGURA 10	– Função de ativação gaussiana	31
FIGURA 11	– Função de ativação linear	31
FIGURA 12	– Exemplo de rede feedforward de camada única	32
FIGURA 13	– Exemplo de rede feedforward de múltiplas camadas	33
FIGURA 14	– Exemplo de rede recorrente	34
FIGURA 15	– Exemplo de rede com estrutura reticulada	34
FIGURA 16	– Objetos linearmente separáveis	35
FIGURA 17	– Problemas não linearmente separáveis	36
FIGURA 18	– Diagrama em blocos de aprendizagem supervisionada	37
FIGURA 19	– Diagrama em blocos de aprendizagem não supervisionada	38
FIGURA 20	– Diagrama em blocos de aprendizagem por reforço	38
FIGURA 21	– Treinamento de uma SVM	41
FIGURA 22	– Cálculo da distância d entre os hiperplanos h_1 e h_2	42
FIGURA 23	– Exemplo de transformação realizada em conjunto de dados não linear	43
FIGURA 24	– Perda do ajuste pela margem suave de uma SVM	45
FIGURA 25	– Arquitetura de uma máquina de regressão construída pelo algoritmo de SV	46
FIGURA 26	– Ilustração da regra de parada antecipada baseada na validação	48
FIGURA 27	– Exemplo de fatores envolvidos na variabilidade espacial de acordo com o grau de probabilidade de ocorrência e facilidade de correção	50
FIGURA 28	– Ciclo da Agricultura de precisão	52
FIGURA 29	– Comparação dos valores reais e estimados da cultura do feijão pela melhor rede neural encontrada	56
FIGURA 30	– Áreas experimentais	58
FIGURA 31	– Grades de amostragens dos talhões 1 e 2	59
FIGURA 32	– SL-31 - Moinho de facas tipo Willey	60
FIGURA 33	– Interface do software JavaNNS	62
FIGURA 34	– Arquivo de entrada no software SNNS e JavaNNS para treinamento de uma rede neural	64
FIGURA 35	– Formato de um arquivo de entrada no software WEKA	64
FIGURA 36	– Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da primeira coleta de nutrientes da folha	72
FIGURA 37	– Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da segunda coleta de nutrientes da folha	72

FIGURA 38 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da terceira coleta de nutrientes da folha	73
FIGURA 39 – Gráfico de comparação dos erros quadráticos médios de treinamento e validação até o término da execução da RNA treinada com os nutrientes da primeira coleta da folha da soja	75
FIGURA 40 – Ilustração do melhor modelo de RNA com os sinais de entrada, os pesos sinápticos, o limiar de ativação, potencial de ativação e função de ativação, utilizando os dados da primeira coleta da folha da soja	76
FIGURA 41 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da segunda coleta de nutrientes da folha na área A ..	78
FIGURA 42 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da primeira coleta de nutrientes da folha	81
FIGURA 43 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da segunda coleta de nutrientes da folha	81
FIGURA 44 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da terceira coleta de nutrientes da folha	81

LISTA DE TABELAS

TABELA 1	– Funções kernel mais utilizadas	44
TABELA 2	– Quantidade absorvida e concentração de nutrientes pela cultura da soja ...	55
TABELA 3	– Concentrações de nutrientes utilizadas para interpretação dos resultados de análise foliar da soja no início do florescimento	55
TABELA 4	– Estatística descritiva do Nitrogênio separada por coletas e áreas.	67
TABELA 5	– Estatística descritiva do Potássio separada por coletas e áreas.	68
TABELA 6	– Estatística descritiva do Fósforo separada por coletas e áreas.	68
TABELA 7	– Estatística descritiva do Magnésio separada por coletas e áreas.	69
TABELA 8	– Estatística descritiva do Cálcio separada por coletas e áreas.	69
TABELA 9	– Estatística descritiva da produtividade da soja separada por áreas.	70
TABELA 10	– Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (pontos amostrais aleatórios para conjunto de treinamento, validação e teste em cada coleta).	71
TABELA 11	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha.	73
TABELA 12	– Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor RNA de cada coleta dos nutrientes da folha da soja.	74
TABELA 13	– Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	75
TABELA 14	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta (utilizando os mesmos pontos amostrais para treinamento, validação e teste para as três coletas).	77
TABELA 15	– Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja para a área A e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	77
TABELA 16	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha somente na área A.	77
TABELA 17	– Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja para a área B e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente	

	de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	78
TABELA 18	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha somente na área B.	79
TABELA 19	– Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja (pontos amostrais aleatórios para conjunto de treinamento, validação e teste em cada coleta).	80
TABELA 20	– Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja.	82
TABELA 21	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta dos nutrientes da folha.	82
TABELA 22	– Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	83
TABELA 23	– Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja somente na área A(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	83
TABELA 24	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta somente da área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	83
TABELA 25	– Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja da área A.	84
TABELA 26	– Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja somente na área B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	84
TABELA 27	– Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta somente da área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	85
TABELA 28	– Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja da área B.	85
TABELA 29	– Atributos selecionados pela técnica CfsSubsetEval para as coletas 1, 2 e 3 das áreas A e B juntas.	87
TABELA 30	– Atributos selecionados pela técnica CfsSubsetEval para as coletas 1, 2 e 3 das áreas A e B separadas.	87
TABELA 31	– Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os	

	atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja nas áreas A e B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta) .	88
TABELA 32 –	Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs utilizando somente os atributos indicados pela técnica CfsSubsetEval de cada coleta das áreas A e B.	89
TABELA 33 –	Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta) .	89
TABELA 34 –	Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta) .	90
TABELA 35 –	Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja nas áreas A e B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	90
TABELA 36 –	Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	91
TABELA 37 –	Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).	91

SUMÁRIO

1	INTRODUÇÃO	17
1.1	OBJETIVO GERAL	18
1.2	OBJETIVOS ESPECÍFICOS	18
1.3	JUSTIFICATIVA	18
1.4	ESTRUTURA DO TRABALHO	19
2	REFERENCIAL TEÓRICO	20
2.1	INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	20
2.2	REDES NEURAIS ARTIFICIAIS	22
2.2.1	História das Redes Neurais Artificiais	23
2.2.2	Funcionamento do Sistema Nervoso	24
2.2.3	Neurônio artificial	26
2.2.4	Funções de ativação	27
2.2.5	Arquiteturas de Redes Neurais Artificiais	31
2.2.6	Redes Perceptron e Adaline	34
2.2.7	Perceptron Multicamadas	35
2.2.8	Processos de treinamento e aprendizagem	36
2.2.9	Algoritmo de Treinamento Backpropagation	38
2.3	MÁQUINAS DE VETORES DE SUPORTE	39
2.3.1	Teoria de Aprendizado Estatístico	40
2.3.2	SVMs Lineares	41
2.3.3	SVMs não lineares	43
2.3.4	Funções Kernel	43
2.3.5	SVMs aplicadas em problemas de regressão	44
2.4	VALIDAÇÃO CRUZADA	46
2.5	COMPARATIVO DE RNA E SVM	48
2.6	AGRICULTURA DE PRECISÃO	49
2.6.1	História da Agricultura de Precisão	51
2.6.2	Ciclo da Agricultura de Precisão	51
2.6.3	Componentes do solo e sua influência nas plantas	52
2.6.4	Aplicação da Inteligência Artificial na agricultura	56
3	MATERIAIS E MÉTODOS	58
3.1	LOCALIZAÇÃO DO EXPERIMENTO	58
3.2	MATERIAIS PARA COLETA DA FOLHA DA SOJA E PARA ANÁLISE FOLIAR	59
3.3	TECNOLOGIAS UTILIZADAS	60
3.4	PRÉ-PROCESSAMENTO DOS DADOS	63
3.4.1	Normalização dos dados	63
3.4.2	Modelagem dos dados	64
3.5	CONJUNTO DE DADOS PARA TREINAMENTO E VALIDAÇÃO	65
3.6	ANÁLISE E COMPARAÇÃO DAS TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL	65
3.6.1	Critério de parada do treinamento	66
3.6.2	Análise Estatística	66

4	RESULTADOS E DISCUSSÃO	67
4.1	ANÁLISE EXPLORATÓRIA DOS DADOS	67
4.2	DESEMPENHO DAS REDES NEURAIAS ARTIFICIAIS	70
4.3	DESEMPENHO DAS MÁQUINAS DE SUPORTE DE VETORES	79
4.3.1	Comparação dos resultados obtidos pelas RNAs e SVMs	85
4.4	SELEÇÃO DE ATRIBUTOS	87
4.4.1	Resultado dos melhores modelos de RNAs obtidos	88
4.4.2	Resultado dos melhores modelos de SVMs obtidos	90
4.4.3	Comparação dos resultados obtidos utilizando a técnica CfsSubsetEval	91
5	CONSIDERAÇÕES FINAIS	93
5.1	CONCLUSÃO	93
5.2	TRABALHOS FUTUROS	94
	REFERÊNCIAS	96

1 INTRODUÇÃO

Segundo dados da Organização para a Cooperação e Desenvolvimento Econômico (2008) (OCDE), a população mundial poderá chegar a 8,2 bilhões de habitantes em 2030, o que significa que a área agricultável por habitante diminuirá, sendo fundamental o aumento da produtividade nas áreas existentes para atender a demanda mundial por alimentos.

Dentre as culturas cultivadas, a soja (*Glycine max*) é uma das oleaginosas mais importantes no mundo e um dos principais produtos do agronegócio do Brasil. Segundo Moreira (2015), uma área de aproximadamente 5,24 milhões de hectares foi cultivada pela soja no Brasil na última safra (2015/2016).

A grande produção da soja se deve a sua maior rentabilidade em relação a outras culturas que são plantadas na mesma época do ano, pois como a cultivar é muito utilizada nas indústrias químicas e de alimentos, para consumo humano e animal e como fonte alternativa de biocombustível, seu preço é mais valorizado (SEDIYAMA, 2009).

Para que a cultura da soja tenha o maior rendimento possível na área plantada, é preciso entender quais nutrientes são relevantes no seu desenvolvimento, pois seus teores influenciam no crescimento e podem servir como auxiliares em uma previsão da produtividade.

Segundo Vitti e Trevisan (2000) as plantas absorvem os nutrientes no decorrer do seu desenvolvimento, e, entre os vários nutrientes, os mais absorvidos pela planta da soja são os macronutrientes, como: Nitrogênio (N), Fósforo (P), Potássio (K), Cálcio (Ca), Magnésio (Mg) e Enxofre (S). Neste trabalho serão testadas e aplicadas as variáveis N, P, K, Ca e Mg para obter um modelo computacional de previsão da produtividade da soja e também serão utilizados métodos de seleção de variáveis para utilizar dentre os 5 nutrientes somente os nutrientes que tem algum tipo de contribuição para cada técnica.

Portanto, este trabalho estudou técnicas de Inteligência Artificial (IA) que estimem os valores da produtividade por meio da quantidade de nutrientes presentes na folha da soja. Desta forma, o estudo pode contribuir com a antecipação dos resultados da produtividade dos grãos, permitindo que deficiências de nutrientes na planta possam ser corrigidas com adubação e calagem no solo em um estágio inicial do desenvolvimento, buscando aumentar a produtividade.

1.1 OBJETIVO GERAL

Fazer um estudo comparativo entre técnicas de Inteligência Artificial para regressão, objetivando avaliar qual a melhor técnica para previsão da produtividade da soja, sendo elas: Redes Neurais Artificiais (RNAs) e Máquinas de Suporte de Vetores (SVMs, do inglês – Support Vector Machines) por serem as técnicas mais utilizadas para classificação e regressão com resultados diversificados para cada caso de estudo.

1.2 OBJETIVOS ESPECÍFICOS

Este trabalho será composto pelos seguintes objetivos específicos:

- Modelar computacionalmente o problema de estimar a produtividade de soja, utilizando informações sobre os nutrientes presentes na folha da planta;
- Utilizar técnicas de seleção de variáveis para selecionar as variáveis mais importantes e relevantes a serem aplicadas nas técnicas de Inteligência Artificial;
- Aplicar e avaliar o desempenho de Redes Neurais Artificiais (RNAs) e Máquinas de Vetores de Suporte (SVMs) como modelo computacional na previsão dos dados da produtividade da soja;
- Justificar a escolha da melhor técnica de Inteligência Artificial para os dados a serem utilizados neste trabalho e por meio dos resultados a serem obtidos através de dados estatísticos e da teoria de aprendizado de máquina.

1.3 JUSTIFICATIVA

A Agricultura de Precisão (AP) tem se tornado fundamental para aumentar a eficiência de áreas produtivas, pois possui técnicas que otimizam o manejo das culturas e a utilização de insumos agropecuários, resultando em menores desperdícios de adubação e fertilizantes em

locais desnecessários da área de plantio, diminuindo a degradação ambiental e aumentando a produtividade (MONDO et al., 2012).

Com a AP, a substituição das tomadas de decisões de gestores e de operadores por sistemas inteligentes se tornará uma necessidade, pois a computação e a automação estão tendo grandes avanços tecnológicos, que contribuem para a melhoria das condições de trabalho na agricultura, aumento da produtividade e para a preservação do meio ambiente (MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO, 2011).

Observa-se na literatura que muitas das técnicas de sistemas inteligentes artificiais já estão sendo aplicadas na agricultura de precisão, explorando o poder de aprendizagem da máquina. Soares et al. (2014) e Leal et al. (2015) aplicaram RNAs na estimativa da produção de grãos da cultura do feijão e do milho, respectivamente. Segundo os autores, as RNAs apresentaram o desempenho esperado, sendo eficientes para estimar a produtividade das culturas estudadas. Bocca (2014), utilizou a SVM como ferramenta de aprendizagem de máquina e conseguiu obter um modelo que estimasse a produtividade da cana de açúcar. Rub (2009) utilizou SVMs e RNAs para previsão da produtividade do trigo e como conclusão do seu trabalho, a SVM foi mais apropriada para os tipos de dados utilizados no seu estudo.

Acredita-se que este trabalho possa comprovar que tecnologias de Inteligência Artificial são capazes de prever a produtividade da soja a partir dos nutrientes presentes na folha. Assim, o produtor terá chances de melhorar a produtividade da sua área de cultivo e aumentar os rendimentos da safra por meio de ferramentas que utilizam Inteligência Artificial, apresentadas neste trabalho e por dados obtidos da análise foliar dos nutrientes da soja.

1.4 ESTRUTURA DO TRABALHO

O trabalho segue com um capítulo de referencial teórico sobre as técnicas de Inteligência Artificial e Agricultura de Precisão aplicadas neste estudo. O capítulo 3 contempla os materiais e métodos utilizados para obtenção dos dados experimentais do trabalho, e para a implementação e avaliação do desempenho das técnicas de inteligência artificial. No quarto capítulo são apresentados e discutidos os resultados obtidos na pesquisa. Por fim, as conclusões finais sobre o estudo e sugestões de trabalhos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo é apresentado o que são técnicas de inteligência artificial, e como as técnicas a serem utilizadas neste trabalho serão redes neurais artificiais e máquinas de vetores de suporte, é dado maior ênfase a elas. Este capítulo também é utilizado para mostrar como a prática da agricultura de precisão está se tornando fundamental na agricultura e como os componentes do solo influenciam no desenvolvimento da planta. Por fim, são apresentados os trabalhos correlatos, que envolvem técnicas de inteligência artificial aplicadas na agricultura para predição da produtividade de diferentes culturas.

2.1 INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

Inteligência artificial é uma área da ciência da computação voltada para o desenvolvimento de sistemas de computadores inteligentes, ou seja, sistemas que exibem características, as quais se relacionam com a inteligência no comportamento do homem, como exemplo: a capacidade de aprender, raciocinar e resolver problemas (FERNANDES, 2003).

O aprendizado de máquina (AM) é um processo de indução de hipóteses a partir de uma experiência. A capacidade de aprendizado é considerada essencial para um comportamento inteligente.

O aprendizado pode ser relacionado com a capacidade de memorizar, observar e explorar situações para aprender fatos, melhorar a capacidade cognitiva/motora e organizar o conhecimento em representações úteis (FACELI et al., 2011).

Segundo Mitchell (1997), AM é a capacidade de melhorar o desempenho na realização de tarefas por meio da experiência.

Algumas aplicações que utilizam AM citados por Faceli et al. (2011):

- Reconhecimento de palavras faladas;
- Predição de taxas de cura de pacientes com diferentes doenças;

- Detecção do uso fraudulento de cartões de crédito;
- Condução de automóveis de forma autônoma em rodovias;
- Ferramentas que jogam gamão e xadrez de forma semelhante a campeões;
- Diagnóstico de câncer por meio da análise de dados de expressão gênica.

AM é utilizada para tarefas de aprendizado preditivas ou descritivas, em que preditivas se referem a previsão de um valor com base nos valores de entrada, a partir de um processo de treinamento. Tarefas descritivas descrevem ou exploram um conjunto de dados.

A Figura 1, apresenta uma hierarquia de aprendizado. Dentro das tarefas de aprendizado indutivo há o aprendizado supervisionado (preditivo) e o não supervisionado (descritivo). Dentro do aprendizado supervisionado, as tarefas são de classificação ou regressão. No caso do aprendizado não supervisionado, são divididas em agrupamento, onde os dados são agrupados de acordo com sua similaridade, e associação, que consiste em encontrar padrões de associações entre os atributos de um conjunto de dados, além da tarefa de sumarização, a qual encontra uma descrição simples e compacta para um conjunto de dados.



Figura 1 – Hierarquia de aprendizado.

Fonte: Faceli et al. (2011).

Definindo melhor as tarefas supervisionadas, pode-se dizer que a classificação visa identificar a qual classe pertence um determinado registro. O modelo analisa os registros já classificados que são informados pelo aplicador dos dados e a partir destes aprende como classificar um novo registro.

A regressão é similar à classificação, porém os registros utilizados são numéricos, não categóricos. Deste modo, a regressão estima o valor de uma determinada variável analisando e associando características com base nos valores informados de entrada que resultaram em determinadas saídas do conjunto de treinamento.

Este trabalho utilizará técnicas de IA para regressão, pois os dados utilizados serão numéricos e não categóricos (classificação), e as entradas da rede neural artificial e da máquina de suporte de vetores receberão os nutrientes da folha da soja, e a saída reproduz o resultado da produtividade da soja. Portanto, nas Seções 2.2 e 2.3 serão descritas mais detalhadamente as técnicas a serem utilizadas neste trabalho: redes neurais artificiais e máquina de vetores de

suporte.

2.2 REDES NEURAIIS ARTIFICIAIS

A Rede Neural Artificial (RNA) é inspirada na observação de como o cérebro humano processa informações. O cérebro é muito complexo, tem a capacidade de organizar os neurônios (constituintes do cérebro) de diferentes formas para realizar atividades, como processamento de padrões, percepção e controle motor muito mais rapidamente que o mais rápido computador hoje existente, devido a sua não linearidade e capacidade de processar em paralelo.

Segundo Haykin (2001), uma RNA é um processador constituído de unidades de processamento simples, que pode armazenar o conhecimento experimental, assemelhando-se ao cérebro humano por adquirir conhecimento por meio de um processo de aprendizagem e pela conexão entre os neurônios, que possuem pesos sinápticos capazes de armazenar o conhecimento adquirido.

As RNAs consistem basicamente em uma quantidade de nós, cada um dos quais é representado por um neurônio artificial. Estes neurônios são organizados em camadas, e os neurônios de uma camada se conectam com os neurônios das camadas adjacentes. Os neurônios de uma RNA computam funções matemáticas. Suas conexões simulam sinapses biológicas e possuem pesos que ponderam os valores de entrada recebidos por cada neurônio da rede. Estes pesos são ajustados conforme o aprendizado da rede (FACELI et al., 2011).

Portanto, uma rede neural pode ser definida como uma simulação simplificada do cérebro humano, pois se assemelha na estrutura de neurônios interconectados e na capacidade de aprender e codificar o conhecimento.

Algumas vantagens das redes neurais citadas por Haykin (2001) são:

- Não-linearidade: se refere a não linearidade distribuída a rede toda;
- Mapeamento de Entrada-Saída: é o aprendizado de máquina supervisionado em que se ajusta os pesos sinápticos por um conjunto de amostras treinadas, até que a rede neural apresente respostas próximas das respostas reais;
- Adaptabilidade: as redes neurais têm capacidade de adaptar os pesos sinápticos às modificações do ambiente em tempo real;
- Resposta a Evidências: além de uma rede neural classificar padrões, fornece informações quanto ao grau de confiabilidade da decisão tomada;

- Informação Contextual: o conhecimento de uma rede neural é representado pela sua estrutura e estado de ativação;
- Tolerância a Falhas: potencial de uma rede neural ser tolerante a falhas. Seu desempenho degrada suavemente sob condições de operações adversas;
- Analogia Neurobiológica: a motivação das redes neurais vem do cérebro humano, uma prova de que o processamento paralelo tolerante a falhas, rápido e poderoso é possível.

Portanto, com as características mencionadas acima, redes neurais artificiais possuem a grande vantagem de solucionar problemas complexos devido às suas capacidades e propriedades não existentes em outras técnicas tradicionais utilizadas para solucionar problemas.

2.2.1 História das Redes Neurais Artificiais

Um dos primeiros estudos da RNA se devem a McCulloch e Pitts. Em 1943, propuseram um modelo matemático de neurônio artificial em que cada neurônio poderia executar diferentes funções lógicas. As primeiras redes não possuíam a capacidade de aprendizagem. Outros pesquisadores como Norbert Wiener e Von Neumann, realizaram estudos na década de 1940, em que sugeriram que as redes neurais fossem inspiradas no funcionamento de um cérebro (HAYKIN, 2001; FACELI et al., 2011).

Hebb (1949) publicou o livro *The Organization of Behavior*, no qual relata sobre a aprendizagem de neurônios por meio de sinapses. Porém, somente em 1956 houve a primeira tentativa de usar simulação computacional pelos pesquisadores Rochester, Holland, Haibt e Duda, para testar uma rede neural com base na teoria proposta no livro de Hebb (HAYKIN, 2001).

O primeiro neuro-computador *Mark I - Perceptron* foi desenvolvido durante os anos de 1957 e 1958 por Frank Rosenblatt, que idealizou o modelo básico do *Perceptron*. O interesse principal de Rosenblatt era o reconhecimento de padrões. Escreveu um livro sobre neuro-computação intitulado como *Principles of Neurodynamics* (YADAV et al., 2015).

Nos anos de 1960, Widrow e Hoff, desenvolveram a rede *Adaline*, do termo *Adaptive Linear Element*. Mais tarde, propôs-se a *Adaline* múltipla, chamada de *Madaline*, rede fundamentada no algoritmo de aprendizado Least Mean Square (LMS) (SILVA et al., 2010).

Na década de 1970, as pesquisas na área de RNAs foram interrompidas, sendo o

principal fator a publicação do livro *Perceptrons – an introduction to computational geometry* por Minsky e Papert em 1969, em que relatava que a rede *Perceptron* de camada única era limitada a realizar a correta classificação apenas dos padrões de classes não linearmente separáveis (HAYKIN, 2001; SILVA et al., 2010).

Na década de 1980, muitos pesquisadores realizaram estudos que contribuíram para a teoria ao projeto de redes neurais artificiais. Os estudos que se destacaram na época foram a derivação de algoritmos de predição utilizando o gradiente reverso por Werbos em 1974, a implementação da rede *Adaptive Resonance Theory* (ART) por Grossberg em 1980 e em 1982, Hopfield utilizou uma função de energia para entender as conexões sinápticas simétricas das redes, que ficaram conhecidas como redes de Hopfield (SILVA et al., 2010).

A retomada de pesquisas na área de redes neurais artificiais se deu ao final dos anos 80 devido aos trabalhos citados no parágrafo anterior, ao desenvolvimento de computadores com maior capacidade de processamento, à criação de algoritmos de otimização e às descobertas sobre o sistema nervoso biológico.

Em 1986 o algoritmo denominado *Backpropagation* foi criado por Rumelhart, Hinton e Williams. Tal algoritmo motivou muito as pesquisas em redes neurais artificiais, pois permitia ajustar os pesos em uma rede com mais de uma camada, resolvendo o problema de aprendizado do reconhecimento de padrões (SILVA et al., 2010).

Vapnik e seus co-autores criaram uma classe de redes de aprendizagem supervisionada, chamada de máquinas de vetores de suporte (SVMs), em que em 1992 os autores Boser, Guyon e Vapnik a inventaram para ser utilizada no reconhecimento de padrões. Em 1995, Cortes e Vapnik inseriram SVMs para regressão e em 1998 Vapnik as aprimorou para problemas de estimação de densidade (HAYKIN, 2001).

Enfim, a história das redes neurais vem sendo construída até os dias atuais, sempre com novas teorias, projetos e aplicações importantes para muitas áreas do conhecimento.

2.2.2 Funcionamento do Sistema Nervoso

O centro do sistema nervoso é o cérebro, representado pela rede neural biológica, a qual recebe continuamente informação e toma decisões apropriadas. Os receptores de informação convertem estímulos do corpo humano ou do ambiente em impulsos elétricos ou químicos, os quais são transmitidos para a rede neural. Estes impulsos são convertidos pelos atuadores em respostas como saída do sistema nervoso (HAYKIN, 2001).

Os principais componentes de um neurônio biológico são: dendritos, corpo celular e axônio e podem ser visualizados na Figura 2. Os dendritos são prolongamentos dos neurônios responsáveis por receber impulsos nervosos de outros neurônios ou do ambiente e conduzi-los ao corpo celular. Dependendo da intensidade e frequência dos estímulos recebidos, o corpo celular gera um novo impulso e envia-o para o axônio. O axônio é um prolongamento dos neurônios, responsável por conduzir os impulsos elétricos gerados no corpo celular para outro local (geralmente dendritos de outros neurônios - neurônios pós-sinápticos). O ponto de contato entre a terminação do axônio de um neurônio e o dendrito de outro neurônio é chamado de sinapse. Portanto, as sinapses fazem a interação entre os neurônios, podendo ser excitatórias ou inibitórias (RUSSELL; NORVIG, 2009).

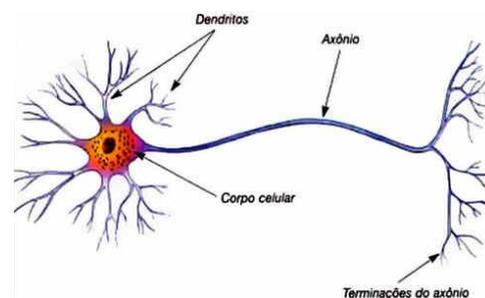


Figura 2 – Neurônio biológico simplificado.

Fonte: Machado (2003).

A Figura 3 representa a hierarquia de organização do cérebro humano. Dos níveis apresentados, o nível de sinapses é o mais fundamental, depende de moléculas e íons para sua ação. O nível seguinte a sinapses é o dos microcircuitos neurais que são um agrupamento de sinapses organizados em padrões de conectividades. Os microcircuitos neurais formam subunidades dendritais dentro das árvores dendritais dos neurônios individuais. O próximo nível é composto pelos circuitos locais constituídos por neurônios com propriedades similares ou diferentes. Estes agrupamentos são seguidos de inter-regionais constituídos por caminhos, colunas e mapas topográficos. Os mapas topográficos são organizados para responder à informação incidente. Os mapas topográficos e outros circuitos inter-regionais são fundamentais para transmitir comportamentos no sistema nervoso central (HAYKIN, 2001).

As redes neurais artificiais ainda estão longe de terem uma estrutura de organização semelhante a do cérebro. Porém, pesquisadores têm estudado e trazido avanços importantes para alcançar gradativamente uma hierarquia similar a apresentada na Figura 3.

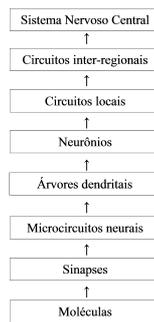


Figura 3 – Organização estrutural dos níveis no cérebro.

Fonte: Haykin (2001).

2.2.3 Neurônio artificial

Neurônio artificial é uma unidade de processamento fundamental para o funcionamento de uma rede neural. É constituído de sete elementos básicos segundo SILVA et al. (2010):

- Sinais de entrada - x_j :
São valores assumidos pelas variáveis a serem utilizadas para treinamento;
- Pesos sinápticos - w_{ki} :
São os pesos que irão ponderar cada um dos valores de entrada. O peso sináptico de um neurônio artificial pode assumir valores negativos ou positivos;
- Combinador linear - Σ :
Soma os sinais de entrada, ponderados pelas respectivas sinapses do neurônio, a fim de produzir um valor de potencial de ativação;
- Limiar de ativação - b_k :
Também chamado de *bias*, aumenta ou diminui a entrada conforme o patamar apropriado para que o resultado produzido pelo combinador linear possa gerar um valor dentro dos limites da função de ativação;
- Potencial de ativação - v_k :
Resultado da diferença do valor produzido entre o combinador linear e o limiar de ativação. Se tal valor ≥ 0 , então o neurônio produz potencial excitatório; caso seja < 0 , o potencial será inibitório;
- Função de ativação - $\varphi(\cdot)$:
Limita a saída do neurônio dentro de um intervalo, geralmente $[0, 1]$ ou $[-1, 1]$;

- Sinal de saída - y_k :

Valor final produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada. Seu valor pode ser utilizado por outros neurônios se houver interligação sequencial.

A Figura 4 representa um modelo de neurônio artificial.

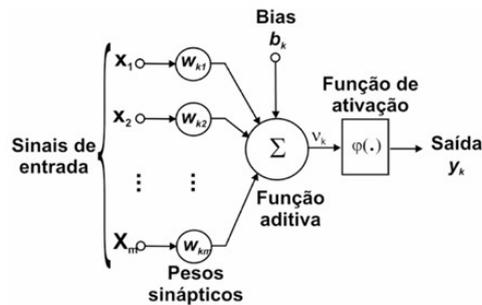


Figura 4 – Modelo de neurônio artificial.

Fonte: Filho et al. (2005).

Em termos matemáticos, um neurônio artificial é representado pelas Equações 1 e 2 (HAYKIN, 2001).

$$v_k = \sum_{j=1}^m w_{kj}x_j \quad (1)$$

$$y_k = \varphi(u_k + b_k) \quad (2)$$

2.2.4 Funções de ativação

As funções de ativação podem ser divididas em funções parcialmente diferenciáveis e totalmente diferenciáveis. O objetivo da função de ativação é limitar a saída do neurônio. Funções de ativação parcialmente e totalmente diferenciáveis são explicadas a seguir, conforme SILVA et al. (2010).

Funções de ativação parcialmente diferenciáveis

Caracterizam-se por possuírem pontos sem derivadas de primeira ordem. As 3 principais funções são: função degrau, função degrau bipolar e função rampa simétrica.

- Função degrau (*Heavyside / hard limiter*)

O resultado da função degrau terá valores unitários positivos quando o potencial de ativação for maior ou igual a zero, caso contrário, terá valor nulo. Representada na Figura 5.

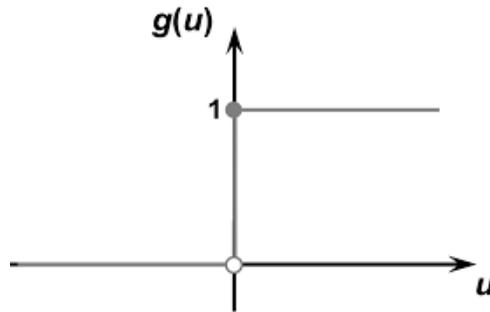


Figura 5 – Função de ativação degrau.

Fonte: SILVA et al. (2010).

- Função degrau bipolar ou função sinal (*symetric hard limiter*)

O resultado desta função obterá valores unitários positivos quando o potencial de ativação for maior que zero, nulo quando o potencial for nulo e valores unitários negativos quando for menor que zero. Ilustrada na Figura 6.

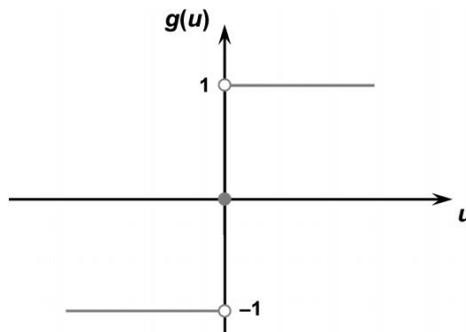


Figura 6 – Função de ativação degrau bipolar.

Fonte: SILVA et al. (2010).

Em classificação de padrões, esta função assume 1 para o potencial de ativação nulo, ou a saída permanece inalterada quando o potencial é nulo.

- Função rampa simétrica

Os valores retornados são os valores dos potenciais de ativação quando estão definidos no intervalo $[-a, a]$, como mostra a Equação 3 e representado pela Figura 7.

$$\begin{cases} a, & \text{se } u > a \\ u, & \text{se } -a \leq u \leq a \\ -a, & \text{se } u < -a \end{cases} \quad (3)$$

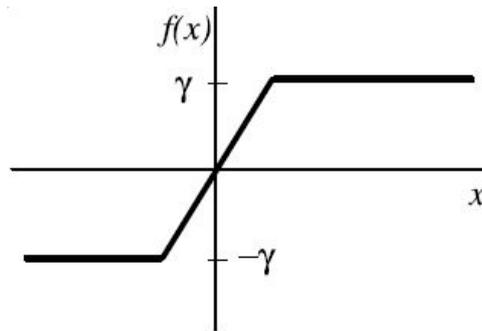


Figura 7 – Função de ativação rampa simétrica.

Fonte: SILVA et al. (2010).

Funções de ativação totalmente diferenciáveis

São funções cujas derivadas de primeira ordem existem e são conhecidas em todos os pontos de seu domínio. As quatro funções que representam este grupo são a função logística, tangente hiperbólica, gaussiana e a função linear.

- **Função logística**

O resultado assumido pela função logística (Figura 8), será sempre entre valores reais de zero a um, expressa na Equação 4.

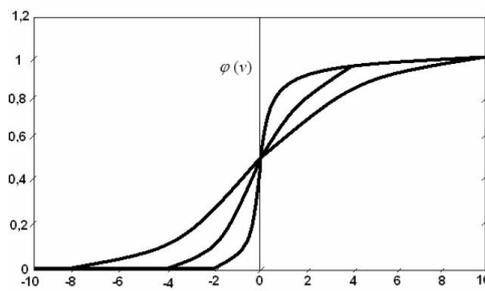


Figura 8 – Função de ativação logística.

Fonte: Cardon e Muller (1994).

$$g(u) = \frac{1}{1 + e^{-\beta \cdot u}} \quad (4)$$

Em que, β é uma constante referente ao nível de inclinação da função. A Figura 8, mostra a variação da função em relação a inclinação β . Quando o parâmetro de inclinação for muito alto (próximo ao infinito), a função de ativação pode ser similar à função degrau. Porém, a função degrau só será 0 ou 1, enquanto a função logística pode ser qualquer valor real entre 0 e 1.

- Função tangente hiperbólica

Os resultados da função tangente hiperbólica sempre estarão no intervalo de números reais de -1 a 1. Assim como ocorre na função logística, quanto maior for o parâmetro de inclinação, mais inclinada será a função tangente hiperbólica (Figura 9). Também, pode se assemelhar a função degrau quando seu valor de inclinação é muito alto. As funções de logística e tangente hiperbólica, são tratadas também de funções sigmóides.

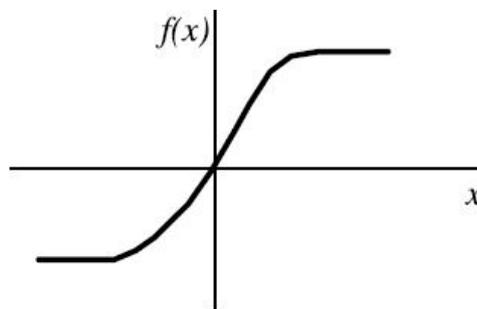


Figura 9 – Função de ativação tangente hiperbólica.

Fonte: Cardon e Muller (1994).

- Função gaussiana

A função gaussiana é expressa matematicamente pela Equação 5:

$$g(u) = \frac{1}{1 + e^{-\frac{(u-c)^2}{2\sigma^2}}} \quad (5)$$

Onde c define o centro e σ representa o desvio padrão da função gaussiana. O valor do desvio padrão possui relação direta com o ponto de inflexão da função. A função está representada na Figura 10.

- Função linear

Também chamada de função identidade, possui resultados produzidos pela saída iguais aos valores do potencial de ativação, ou seja, $f(x) = x$, como mostra a Figura 11.

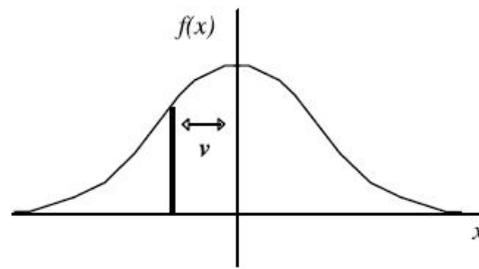


Figura 10 – Função de ativação gaussiana.

Fonte: Cardon e Muller (1994).

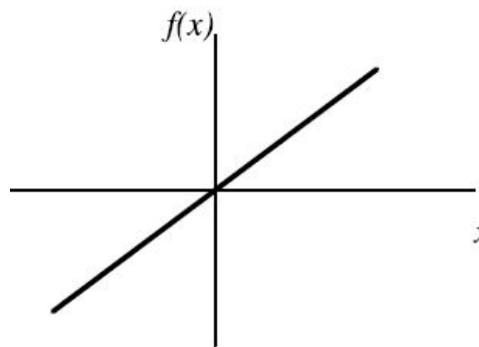


Figura 11 – Função de ativação linear.

Fonte: Cardon e Muller (1994).

2.2.5 Arquiteturas de Redes Neurais Artificiais

A arquitetura de uma RNA define a forma com que os seus neurônios se conectam uns aos outros. Estes arranjos são estruturados por meio do direcionamento das conexões sinápticas dos neurônios. O treinamento de uma determinada arquitetura possui passos específicos para ajustar os pesos e os limiares de seus neurônios. Esse processo de ajuste, é conhecido como algoritmo de aprendizagem, onde sintoniza-se a rede para que as respostas saiam próximas aos valores desejados. Uma rede neural pode ser dividida em três camadas:

- Camada de entrada: responsável por receber os dados advindos do meio externo, geralmente normalizados em relação às faixas de variação produzidas pelas funções de ativação;
- Camadas escondidas: chamadas também de intermediárias, ocultas ou invisíveis. São compostas por neurônios que possuem a capacidade de extrair as características necessárias para o processo de aprendizagem;

- Camada de saída: camada responsável pela produção e apresentação dos resultados finais da rede, advindos de processos realizados nas camadas anteriores.

As principais arquiteturas de RNAs são: redes *feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas (SILVA et al., 2010).

Arquitetura *feedforward* de camada única

Uma RNA com arquitetura *feedforward* de camada única possui uma camada de entrada e uma única camada de neurônios de saída. O fluxo de informações segue a mesma direção durante todo o processo, ou seja, sempre da camada de entrada para a camada de saída (Figura 12). A quantidade de saídas da rede neural nesta arquitetura, sempre coincidirá com a quantidade de neurônios da camada de entrada.

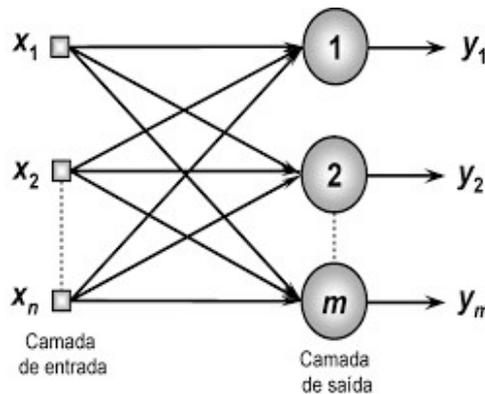


Figura 12 – Exemplo de rede feedforward de camada única.

Fonte: SILVA et al. (2010).

Essa arquitetura é utilizada para problemas de reconhecimento de padrões e filtragem linear. Como exemplos destas arquiteturas existem *Perceptron* e *Adaline* baseadas nas regras de Hebb e Delta, respectivamente (SILVA et al., 2010; HAYKIN, 2001).

Arquitetura *feedforward* de múltiplas camadas

A arquitetura *feedforward* de múltiplas camadas é semelhante a anterior, porém pode ter uma ou mais camadas escondidas de neurônios. São aplicadas para aproximação de funções, classificação de padrões, identificação de sistemas, otimização robótica, ao controle de processos.

A Figura 13 representa uma rede com uma camada de entrada composta de n sinais, duas camadas escondidas de n_1 e n_2 neurônios, e uma camada neural de saída com m neurônios. De acordo com Faceli et al. (2011), uma rede com múltiplas camadas pode ser classificada em:

- Completamente conectada: quando os neurônios da rede estão todos conectados a todos os neurônios da camada anterior e/ou seguinte;
- Parcialmente conectada: quando os neurônios estão conectados a apenas alguns neurônios da camada anterior e/ou seguinte;
- Localmente conectada: são redes parcialmente conectadas em que os neurônios que se conectam a um outro neurônio são todos de uma região bem definida.

Os principais tipos de rede com arquitetura *feedforward* de múltiplas camadas são o *Perceptron* Multicamadas (*Multilayer Perceptron* - MLP) e as redes de base radial (*Radial Basis Function* – RBF) baseados na regra delta generalizada e delta/competitiva.

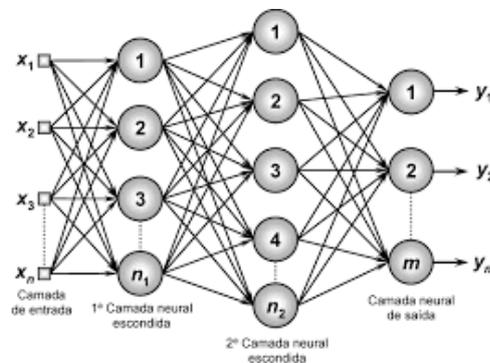


Figura 13 – Exemplo de rede feedforward de múltiplas camadas.

Fonte: SILVA et al. (2010).

Arquitetura recorrente ou realimentada

A saída de um neurônio é realimentada como sinais de entrada para outros neurônios. A realimentação permite que estas redes possam ser utilizadas em sistemas variantes em relação ao tempo, como previsão de séries temporais, otimização e identificação de sistemas, controle de processos. Exemplos de redes realimentadas são a rede de Hopfield e Perceptron multicamadas, cujas redes são baseadas nas regras de minimização de funções de energia e na regra delta generalizada. A Figura 14 mostra uma camada intermediária sendo realimentada por um neurônio da camada de saída. Com esta arquitetura, as saídas futuras levam em consideração, o resultado das saídas anteriores (SILVA et al., 2010; HAYKIN, 2001).

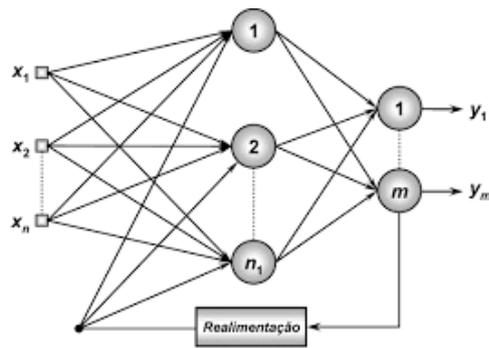


Figura 14 – Exemplo de rede recorrente.

Fonte: SILVA et al. (2010).

Arquitetura em estrutura reticulada

Dispõe os neurônios conforme a necessidade de ajuste de seus pesos e limiares. Aplicada para problemas de agrupamento (*clustering*), reconhecimento de padrões, otimização de sistemas, grafos. Um exemplo de rede que utiliza esta arquitetura é a rede Kohonen (Figura 15), em que seus neurônios estão dispostos no espaço bidimensional, com o propósito de realizar o treinamento por meio de um processo competitivo. Nesta topologia, os sinais de entrada são inseridos em todos os neurônios da rede (SILVA et al., 2010; HAYKIN, 2001).

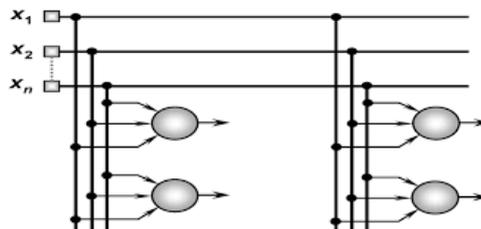


Figura 15 – Exemplo de rede com estrutura reticulada.

Fonte: SILVA et al. (2010).

2.2.6 Redes Perceptron e Adaline

As redes *Perceptron* e *Adaline* pertencem a arquitetura *feedforward* de camada única. A rede *Perceptron* é treinada por um algoritmo supervisionado de correção de erro e usa a função de ativação do tipo limiar. Durante seu treinamento, para uma entrada, os pesos são ajustados conforme a Equação 6 :

$$w_j(t+1) = w_j(t) + \eta x_i^j (y_i - f(x_i)) \quad (6)$$

Em que $w_j(t)$ é o peso da j -ésima conexão de entrada no tempo t , η é uma taxa de aprendizado, x_i^j é o valor do j -ésimo atributo do vetor de entrada, $f(x_i)$ é a saída produzida pela rede no tempo t e y_i é a saída desejada para a variável de x_i . A diferença da rede *Perceptron* para *Adaline* é que a *Adaline* utiliza a função de ativação linear, assim leva a magnitude do erro em consideração na hora de ajustar os pesos da rede. Para isso, a rede *Adaline* utiliza uma regra chamada Delta, cujo nome é da diferença entre os valores da saída desejada e da saída produzida (FACELI et al., 2011). *Adaline* é utilizada para problemas de regressão, e *Perceptron* para problemas de classificação. Tanto *Perceptron* quanto *Adaline*, só classificam objetos linearmente separáveis como representado na Figura 16.

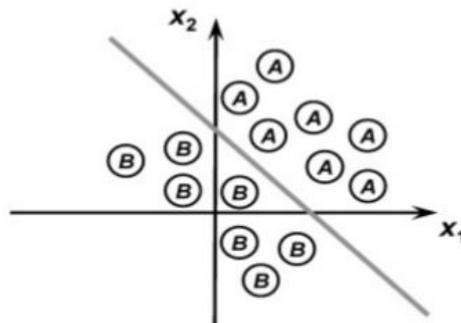


Figura 16 – Objetos linearmente separáveis.

Fonte: SILVA et al. (2010).

2.2.7 Perceptron Multicamadas

Os problemas não linearmente separáveis são resolvidos com RNAs *Perceptron* Multicamadas (Figura 17).

Possuem uma ou mais camadas intermediárias, podendo implementar qualquer função contínua. A arquitetura mais utilizada para uma rede *Perceptron* Multicamadas é a completamente conectada. A primeira camada, aprende uma função em cada neurônio que define um hiperplano com uma divisão de duas partes. Cada neurônio da camada seguinte, combina sua função com a função do neurônio da camada anterior, formando regiões convexas

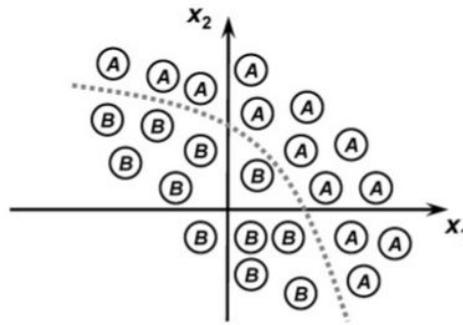


Figura 17 – Problemas não linearmente separáveis.

Fonte: SILVA et al. (2010).

e assim sucessivamente, sendo uma função mais complexa a cada camada da rede. É utilizado para problemas de classificação e regressão.

2.2.8 Processos de treinamento e aprendizagem

Uma rede neural artificial se destaca pela sua capacidade de aprender a partir do seu ambiente e melhorar o seu desempenho à medida que treina, pois aprende por meio de um processo interativo de ajustes aplicados aos pesos sinápticos e níveis do *bias*. Uma definição de aprendizado de máquina é explanada em uma sequência de eventos por Haykin (2001):

1. A rede neural é estimulada por um ambiente;
2. A rede neural sofre modificações nos seus parâmetros como resultado da estimulação;
3. A rede neural responde ao ambiente de forma diferente, devido às modificações ocorridas na sua estrutura interna.

As amostras a serem aplicadas em uma rede neural artificial, geralmente são divididas em dois subconjuntos: subconjunto para treinamento e outro para teste. O subconjunto para treinamento possui de 60 a 90% das amostras do conjunto total. Este conjunto de treinamento é utilizado no processo de aprendizagem da rede. Os outros 10 a 40% restantes, são utilizados para teste, onde é verificado se a rede pode ser validada (SILVA et al., 2010). O processo de treinamento em que se ajusta os pesos sinápticos e limiares dos neurônios é denominado de época de treinamento. Serão explicados nas próximas subseções três tipos de treinamento de redes neurais.

Aprendizagem supervisionada

É um caso de inferência indutiva, em que as saídas desejadas ao sistema investigado são conhecidas. Neste caso, é utilizada uma tabela atributos/valores, em que há valor da saída desejada com as respectivas entradas. A aprendizagem supervisionada funciona como se houvesse um professor ensinando para a rede se a saída obtida é desejada ou não e no caso de não ser a desejada, o supervisor tenta ajudá-la a melhorar. Deste modo, os pesos sinápticos e limiares são continuamente ajustados mediante a aplicação de ações comparativas, de acordo com o algoritmo de aprendizagem, as quais supervisionam o erro em relação as respostas produzidas àquelas desejadas, sendo este erro usado para realizar os ajustes. A rede só é considerada treinada, quando os valores de defasagem forem considerados aceitáveis no conjunto de dados (SILVA et al., 2010; HAYKIN, 2001). A Figura 18 representa esta forma de aprendizagem.

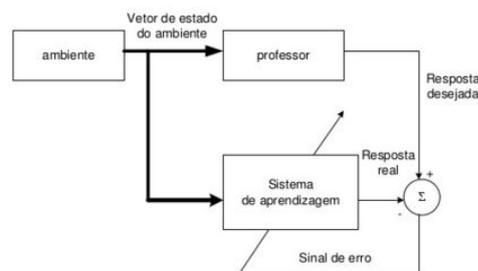


Figura 18 – Diagrama em blocos de aprendizagem supervisionada.

Fonte: Silva (2001).

Aprendizagem não supervisionada

Neste tipo de aprendizagem, as saídas desejadas não existem. Consequentemente, a rede neural tem que se auto organizar para entender conjuntos com similaridades. O projetista da rede pode determinar a quantidade máxima de agrupamentos (clusters), levando-se em consideração seu conhecimento a respeito do problema a ser tratado (Figura 19).

Aprendizagem por reforço

É considerada uma variação de técnicas que utilizam aprendizagem supervisionada. Os algoritmos utilizados na aprendizagem por reforço ajustam os parâmetros internos dos neurônios, baseando-se em características quantitativas ou qualitativas advindas da interação com o ambiente, as quais medem o desempenho do aprendiz. É um processo de tentativa e erro, pois o tipo de resposta obtida de uma determinada entrada é se esta é satisfatória ou não (SILVA et al., 2010). A Figura 20 mostra um diagrama do funcionamento da aprendizagem por reforço.



Figura 19 – Diagrama em blocos de aprendizagem não supervisionada.

Fonte: Silva (2001).

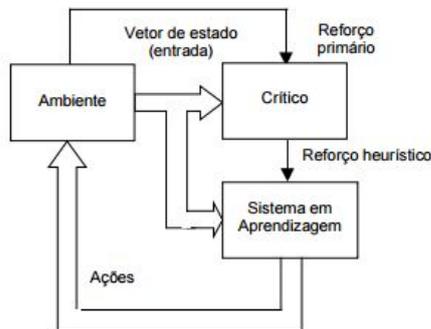


Figura 20 – Diagrama em blocos de aprendizagem por reforço.

Fonte: Silva (2001).

2.2.9 Algoritmo de Treinamento Backpropagation

Este algoritmo se baseia na regra delta utilizada na rede *Adaline*, também conhecida como regra delta generalizada. Possui duas fases, a fase para frente (*forward*) e para trás (*backward*). Na fase *forward*, cada neurônio da primeira camada intermediária, recebe os valores já ponderados pelas conexões de entrada e, a partir deste, produz os valores de saída que serão utilizados como entrada para a sua camada seguinte, assim sucessivamente até que os neurônios da camada de saída produzam cada um o seu valor de saída. Ao fim deste processo, é analisado o erro, que é a diferença das saídas produzidas e dos valores desejados (FACELI et al., 2011).

Este erro é utilizado na fase *backward*, onde cada neurônio ajusta os pesos da rede, desde a camada de saída até a primeira camada intermediária. O algoritmo *Backpropagation*

faz com que um erro de uma determinada camada intermediária seja estimado com a soma dos erros dos neurônios da camada seguinte (MITCHELL, 1997; SILVA et al., 2010).

A taxa de aprendizado tem forte influência na rede, pois uma pequena taxa de aprendizado pode gerar muitos ciclos desnecessários para induzir um bom modelo. Já uma taxa elevada, pode provocar oscilações que dificultam a criação de uma rede boa. Para este problema, criou-se o termo *momentum*, o qual quantifica o grau de importância da variação de peso do ciclo anterior ao atual (FACELI et al., 2011; MITCHELL, 1997).

Podem ser utilizados critérios de paradas no treinamento, como um número máximo de ciclos ou uma taxa máxima de erro. O conjunto de validação, é utilizado para reduzir a ocorrência de *overfitting*, o qual ocorre quando a taxa de erro começa a subir, sendo um indício de que a rede está decorando os dados. Então utiliza-se também, o critério de parada quando a taxa de erro para o conjunto de validação começa a subir (FACELI et al., 2011).

2.3 MÁQUINAS DE VETORES DE SUPORTE

As máquinas de vetores de suporte são redes *feedforward*, embasadas pela teoria de aprendizado estatístico que foram desenvolvidas por Vapnik em 1995 a partir de estudos iniciados no livro *The Nature of Statistical Learning Theory* por Vapnik e Chervonenkis (1971), no trabalho de Boser, Guyon e Vapnik (1992), no artigo *Support-Vector Networks* escrito por Cortes e Vapnik (1995), e em muitas outras obras publicadas (HAYKIN, 2001).

SVMs podem ser usadas para resolver problemas de classificação de padrões e regressão linear. Mas, SVM surgiu primeiramente como classificadora, a qual classifica com boa generalização e com capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu (LORENA; CARVAHO, 2003).

No contexto de classificação de padrões, a SVM, constrói um hiperplano de decisão de tal forma que exemplos positivos e negativos são separados por uma margem máxima. A SVM é implementada pelo método de minimização estrutural de risco, em que a taxa de uma máquina de aprendizagem sobre dados de teste é limitada pela soma da taxa de erro de treinamento e por um outro termo dependente da dimensão de Vapnik. Será mais detalhado adiante o funcionamento de SVM como classificadora e regressora.

As SVMs se caracterizam por possuírem boa capacidade de generalização, ou seja, eficiência em classificar um conjunto de dados que não pertencem ao conjunto de treinamento.

Sua teoria é bem definida dentro da Matemática e Estatística. Outra característica de SVMs é sua robustez como objetos de grandes dimensões, como imagens. Também, possui convexidade na função objetivo, o que implica na otimização de uma função quadrática, que possui apenas um mínimo global (SCHÖLKOPF; SMOLA, 2001).

Os autores Faceli et al. (2011) citam como exemplo o destaque de SVMs como classificadores no trabalho de Joachims (2002), que aplicou SVM na categorização de textos. Também, cita os estudos de Schölkopf et al., (2003) e Noble (2004) na Bioinformática utilizando SVMs. O autor Burges (1998) dá outros exemplos no caso de classificação como reconhecimento de objetos (BLANZ et al., 1996), identificação do locutor (SCHIMIDT, 1996), detecção de rosto em imagens (OSUNA, FREUND; GIROSI, 1997a) e no uso de SVMs como regressoras. SVMs foram utilizadas para comparar testes de previsão de séries temporais de referência (MULLER et al., 1997; MUKHERJEE, OSUNA; GIROSI, 1997).

2.3.1 Teoria de Aprendizado Estatístico

A Teoria de Aprendizado Estatístico (TAE) estabelece condições matemáticas que auxiliam na escolha de um classificador a partir de um conjunto de dados de treinamento. O erro esperado de um classificador, é também chamado de risco, que mede a capacidade de generalização do classificador. A TAE, provê diversos limites de risco esperados de uma função de classificação, os quais podem ser utilizados na escolha de um classificador.

O treinamento de uma máquina de vetores de suporte consiste em encontrar o hiperplano ótimo, isto é, aquele com a distância máxima dos padrões de treinamento mais próximos. Os vetores de suporte são os padrões (mais próximos) a uma distância b do hiperplano. Na Figura 21, três vetores de suporte são mostrados em pontos sólidos (DUDA et al., 2001).

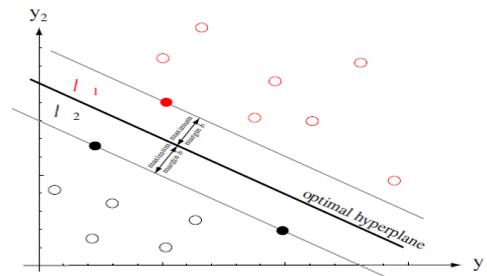


Figura 21 – Treinamento de uma SVM.

Fonte: Duda et al. (2001).

2.3.2 SVMs Lineares

As SVMs lineares são eficazes na classificação de conjuntos de dados linearmente separáveis ou que possuam uma distribuição aproximadamente linear. Utiliza fronteiras lineares para a separação de objetos pertencentes a duas classes, podendo ter a presença de alguns ruídos e *outliers* (LORENA; CARVAHO, 2003). Porém, muitas vezes há dados que não são divididos satisfatoriamente por um hiperplano, assim tornou-se necessário a obtenção de fronteiras não lineares com SVMs, por meio de uma extensão das SVMs lineares.

SVMs com Margens Suaves

Em problemas não linearmente separáveis existe a presença de ruídos e *outliers* nos dados. Para resolver estes problemas, permite-se que alguns dados possam violar a restrição da Equação 7. Para isto, introduz-se variáveis de folga ε_i , para todo $i = 1, \dots, n$. Essas variáveis relaxam as restrições impostas ao problema de otimização primal, que se tornam:

$$y_i(w \cdot x_i + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0, \forall i = 1, \dots, n \quad (7)$$

A aplicação desse procedimento suaviza as margens do classificador, por isto é chamado de SVMs com margens suaves, pois pode ter hiperplanos com objetos entre dois hiperplanos permitindo a ocorrência de alguns erros de classificação (FACELI et al., 2011).

SVMs com Margens Rígidas

As SVMs com margens rígidas definem fronteiras lineares a partir de dados linearmente separáveis. Um conjunto de treinamento que é linearmente separável, separa os objetos em duas classes por um hiperplano. Os classificadores que separam os dados por meio de um hiperplano são denominados lineares. A Equação de um hiperplano é apresentada na

Equação 8 (FACELI et al., 2011).

$$h(x) = u \cdot a + d \quad (8)$$

Onde $u \cdot a$ é o produto escalar entre os vetores u e a , onde u pertence ao conjunto de treinamento e é o vetor normal ao hiperplano descrito, d é um termo “compensador”. O par (u, d) é determinado durante o processo de treinamento do classificador (LORENA; CARVAHO, 2003). A partir de $h(x)$, é possível obter infinitos hiperplanos equivalentes multiplicando u e d por uma mesma constante. O hiperplano canônico é definido como aquele em que u e d são escalados de forma com que os exemplos mais próximos ao hiperplano $u \cdot a + d = 0$ satisfaçam as Equações 9 e 10.

$$|u \cdot a_i + d| = 1 \quad (9)$$

$$|u \cdot a_i + d| = -1 \quad (10)$$

Seja a_1 um ponto no hiperplano $H_1 : u \cdot a + d = +1$ e a_2 um ponto no hiperplano $H_2 : u \cdot a + d = -1$, ilustrado na Figura 22 . É possível obter a distância entre os hiperplanos H_1 e H_2 projetando $a_1 - a_2$ na direção de u , perpendicular ao hiperplano separador $u \cdot a + d = 0$, conforme Equação 11 (LORENA; CARVALHO, 2007).

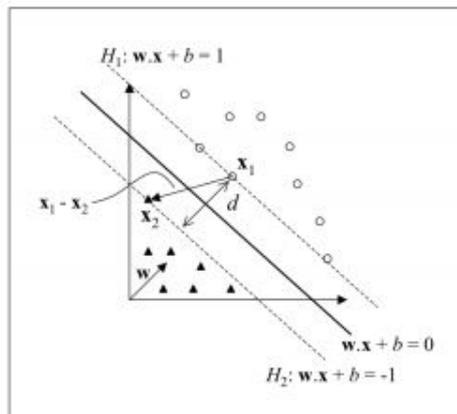


Figura 22 – Cálculo da distância d entre os hiperplanos h_1 e h_2 .

Fonte: Lorena e Carvalho (2007).

$$\frac{2(a_1 + a_2)}{\|u\| \|a_1 - a_2\|} \quad (11)$$

2.3.3 SVMs não lineares

Na Figura 23, os objetos estão em R^2 e não podem ser divididos linearmente, então seria mais adequado dividi-los em outra dimensão (R^3) para se tornarem linearmente separáveis.

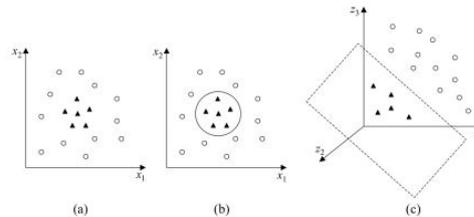


Figura 23 – Exemplo de transformação realizada em conjunto de dados não linear. (a) Conjunto de dados. (b) Fronteira para duas dimensões. (c) Fronteira para três dimensões.

Fonte: Lorena e Carvalho (2007).

Seja $\Phi : Z \rightarrow \gamma$ um mapeamento, em que Z é o espaço de entradas e γ denota o espaço de características. A escolha de Φ faz com que o conjunto de treinamento mapeado em γ possa ser separado por uma SVM linear. Para este mapeamento de espaço há duas condições baseadas no teorema de Cover (HAYKIN, 2001): A primeira é que a transformação seja não linear e a segunda condição é que a dimensão do espaço de características seja suficientemente alta. Portanto, para casos como o exemplo da Figura 23, em que duas dimensões não são suficientes para separar os dados linearmente, mapeia-se os objetos para um espaço de maior dimensão e aplica-se a SVM linear sobre o espaço.

2.3.4 Funções Kernel

γ Pode ter dimensão muito alta, dificultando a computação de Φ . A única informação necessária sobre o mapeamento é de como realizar o cálculo de produtos escalares entre os objetos no espaço de características, e pode ser obtida pelo uso de funções denominadas kernels. Portanto, a importância de kernels está na simplicidade de seu cálculo e na sua capacidade de representar espaços abstratos. As funções Φ devem pertencer a um domínio em que seja possível o cálculo de produtos internos. Geralmente faz-se uso das condições estabelecidas pelo Teorema de Mercer em que os kernels devem ser matrizes positivamente definidas, isto

é, a matriz K , em que $K_{ij} = K(x_i, x_j)$ para todo $i, j = 1, \dots, n$ maiores do que zero (MERCER, 1909). Segundo Faceli et al. (2011), Lorena e Carvaho (2003), Haykin (2001), alguns dos kernels mais utilizados são os polinomiais, de RBF ou gaussiano e os sigmoidais (Tabela 1).

Tabela 1 – Funções kernel mais utilizadas.

Tipo de Kernel	Função $K(x_i, x_j)$ correspondente	Comentários
Polinomial	$(x_i^j \cdot x_j + 1)^p$	A potência p deve ser especificada pelo usuário
Gaussiano	$\exp\left(-\frac{1}{2\sigma^2} \ x_i - x_j\ ^2\right)$	A amplitude σ^2 é especificada pelo usuário
Sigmoidal	$\text{Tanh}(\beta_0 x_j \cdot x_j + \beta_1)$	Utilizado somente para alguns valores de β_0 e β_1

Fonte: Lorena e Carvaho (2003).

Segundo Lorena e Carvaho (2003), quase todas formas de mapeamento podem se implementadas por esta função. Além disso, por meio dessa função, define-se Redes Neurais do tipo RBF, em que são apresentadas no livro de Mitchell (1997). Redes Neurais do tipo RBF possuem um número de funções de bases radiais determinados pelo número de *Support Vectors* (SVs) e seus centros determinados pelos valores de multiplicadores de Lagrange (LORENA; CARVAHO, 2003).

2.3.5 SVMs aplicadas em problemas de regressão

A SVM foi desenvolvida inicialmente para reconhecimento de caracteres. Dentro de um período de tempo, se tornou um dos melhores sistemas disponíveis para classificação. Mas também, logo foram obtidas aplicações de previsão na regressão e de séries temporais com excelentes desempenhos relatados por Muller et al., 1997, Drucker et al., 1997, Stitson et al., 1999, Mattera e Haykin 1999 (SMOLA; SCHOLKOPF, 2004). O algoritmo Support Vector Regression (ϵ -SVR) tem como objetivo encontrar uma função $h(x)$ que produza saídas contínuas para os dados de treinamento com um desvio máximo de ϵ de seu rótulo desejado. A função deve ser o mais regular possível, o que significa buscar uma função com um q pequeno, por meio da minimização da norma $\|q\|$.

$$\text{Minimizar } \frac{1}{2} \|q\|^2 \quad (12)$$

$$\text{Com as restrições: } \begin{cases} Y_i - q \cdot X_i - b \leq \varepsilon_i \\ q \cdot X_i + b - Y_i \leq \varepsilon_i \end{cases} \quad (13)$$

Este problema é tratado como um problema de otimização. De acordo com as Equações 12 e 13, procura-se uma função linear que aproxime (x_i, y_i) de treinamento com precisão ε .

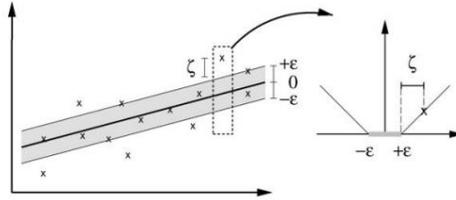


Figura 24 – A perda do ajuste pela margem suave de uma SVM.

Fonte: Lorena e Carvaho (2003).

A Figura 24 representa graficamente o procedimento realizado. Busca-se uma função em que os dados de treinamento fiquem dentro da região entre $+\varepsilon$ e $-\varepsilon$ (parte sombreada da Figura 24). Porém, como em SVMs com margens suaves, algumas variáveis de folga são introduzidas, devido o problema de ter que lidar com pontos fora da área sombreada, permitindo assim, os ruídos e *outliers* nos objetos. Na Figura 24, as variáveis de folga são representadas pelos pontos que estão fora da região entre $-\varepsilon$ e $+\varepsilon$. A Equação utilizada para minimizar a norma e obter o problema de otimização do exemplo dado é representada pela Equação 14 e 15 (SMOLA; SCHOLKOPF, 2004; FACELI et al., 2011).

Minimizar q, b, ξ, ξ^* :

$$\frac{1}{2} \|q\|^2 + C \left(\sum_{i=1}^n \xi_i + \xi^* \right) \quad (14)$$

Com as restrições:

$$\begin{cases} y_i - q \cdot x_i - b \leq \varepsilon_i + \xi \\ w \cdot x_i + b - y_i \leq \varepsilon_i + \xi^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (15)$$

ξ_i, ξ_i^* Representam as variáveis de folga e $C > 0$ é uma constante que determina o *trade-off* entre a regularidade da faixa de $-\varepsilon$ e $+\varepsilon$ e o quanto de desvios são tolerados. Monta-se então, o problema dual usando função de Lagrange, tornando nulo o resultado das derivações parciais e substituindo as expressões resultantes na Equação lagrangiana inicial. Também pode ser usado kernels em casos de regressões não lineares. O problema de otimização final é dado pelas Equações 16 e 17 (SMOLA; SCHOLKOPF, 2004).

Maximizar α_i, α_j^* :

$$\frac{-1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_j^*) \varpi(x_i, x_j) - \varepsilon \sum_{i=1}^n (\alpha_i - \alpha_j^*) + \sum_{i=1}^n y_i (\alpha_i - \alpha_j^*) \quad (16)$$

Com as restrições: $\begin{cases} \sum_{i=1}^n (\alpha_i - \alpha_j^*) = 0 \\ \alpha_i, \alpha_j^* \in [0, C] \end{cases} \quad (17)$

Onde α_i e α_j^* representam as variáveis de Lagrange e ϖ a função kernel, que deve satisfazer as condições de Mercer (conforme explicado para SVMs de classificação, a Tabela 1 também se aplica às SVRs). As variáveis que se encontram dentro da margem entre $-\varepsilon$ e $+\varepsilon$ são nulas.

A Figura 25 contém um gráfico de visão geral sobre as diferentes etapas na fase de regressão de uma SVR.

O padrão de entrada (para os quais uma predição é feita) são mapeados para o espaço de características por um mapa Φ . Então os produtos (*dot product*) são calculados com as imagens dos padrões de treinamento por meio do mapa Φ . Isto corresponde a avaliação de funções de kernel $\varpi(x, x_i)$. Finalmente, os produtos de ponto são somados utilizando os pesos $v_i = \alpha_i - \alpha_i^*$. Isto, mais o termo b constante produz a saída final de predição (SMOLA; SCHOLKOPF, 2004).

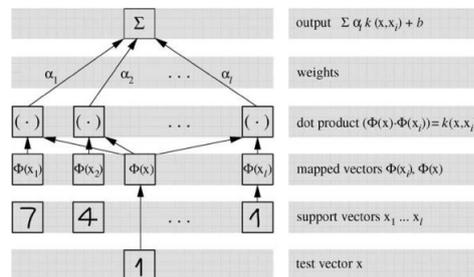


Figura 25 – Arquitetura de uma máquina de regressão construída pelo algoritmo de SV.

Fonte: Smola e Scholkopf (2004).

2.4 VALIDAÇÃO CRUZADA

Uma ferramenta padrão da estatística conhecida para validar o modelo é chamada de validação cruzada, onde se utiliza um conjunto de dados diferente daquele usado no treinamento

para estimar os parâmetros (HAYKIN, 2001). O conjunto de dados disponível para criar uma rede neural artificial é dividido em dois subconjuntos, um dos subconjuntos é usado para treinamento e validação, ou seja, para obter um modelo de rede neural que atinja o objetivo esperado e o outro subconjunto, é usado para o teste deste modelo de rede, onde será testado valores desconhecidos para a rede neural artificial treinada, buscando verificar o quanto o modelo escolhido prediz os dados reais não informados no conjunto de teste. Esta partição dos dados deve ser repetida várias vezes durante o processo de aprendizado. O desempenho de cada topologia candidata é medido pela média dos desempenhos individuais em cada experimento utilizando os dados do subconjunto de teste até encontrar um erro aceitável. A este processo se denomina o termo validação cruzada (SILVA et al., 2010).

No entanto, há a possibilidade do melhor modelo avaliado pela estatística de validação cruzada ter ajustado excessivamente os parâmetros, fazendo com que o modelo decore os dados e não os aprenda. Portanto, para isto pode se utilizar critérios de otimização dos modelos para verificar o momento certo de parar de ajustar os parâmetros e encerrar o treinamento.

Uma heurística que é citada por Haykin (2001) chamado método de treinamento com parada antecipada, busca analisar o erro de treinamento que começa alto e decresce com o aumento do número de épocas em direção a um mínimo local na superfície do erro e analisa não somente a curva de aprendizagem de treinamento, mas também a curva do erro de validação da rede. Assim, pode se perceber quando é o melhor momento para encerrar o treinamento da RNA, para não treiná-la excessivamente ao ponto de decorar os dados (*overfitting*) em vez de aprendê-los.

Portanto, para esta heurística utiliza-se uma divisão dos dados em três conjuntos para verificar o início do excesso de treinamento. Divide-se os dados em um conjunto de treinamento, um conjunto de validação e um conjunto para teste. O subconjunto de treinamento é utilizado para treinar a rede. Após um período de treinamento, em que os pesos sinápticos e os níveis de bias da rede neural são fixos, é medido o erro de validação para cada ponto do subconjunto de validação. Quando a fase de validação é completada, o treinamento é reiniciado para um novo período, e o processo é repetido.

A Figura 26 mostra a curva de aprendizagem de validação e treinamento, e nota-se que a curva de aprendizagem de validação decresce monotonamente até um mínimo local e então começa a crescer conforme o treinamento continua.

O método de parada antecipada sugere que o ponto mínimo na curva de aprendizagem de validação seja o critério de parada do treinamento da rede neural artificial e o melhor modelo escolhido dentre as RNAs treinadas, seja aquele em que obteve-se o menor erro de validação.

Há muitas variantes de validação cruzada, como a validação cruzada múltipla, quando

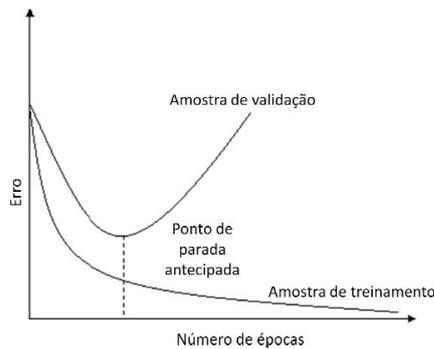


Figura 26 – Instrução do método de treinamento com parada antecipada baseada na validação.

Fonte: Menezes (2014).

se há pequenas quantidades de dados. Divide-se o conjunto disponível de N exemplos em K subconjuntos. O modelo é treinado com todos os subconjuntos, exceto um, e o erro de validação é obtido por meio deste subconjunto não utilizado no treinamento. Este procedimento é repetido K vezes, utilizando em cada uma um subconjunto de validação diferente. O desempenho é medido pela média do erro quadrado obtida pelas K tentativas do experimento. Porém, esta técnica possui a desvantagem de realizar uma quantidade excessiva de cálculos, pois o modelo deve ser treinado muitas vezes (HAYKIN, 2001).

2.5 COMPARATIVO DE RNA E SVM

As duas técnicas são consideradas “caixas-pretas”, ou seja, o conhecimento extraído por elas é codificado em equações de difícil interpretação.

As RNAs mais empregadas na prática, segundo Faceli et al. (2011), são as redes MLP, e o *back-propagation* como sendo o algoritmo de treinamento mais popular. Este algoritmo é baseado no método de otimização gradiente descendente e busca por minimizar o erro quadrático entre os valores desejados e os obtidos pela rede, enquanto SVMs, são baseadas a um problema de otimização quadrático em função dos dados de treinamento, buscando maximizar a margem de separação entre as classes, sendo uma única solução global, já a função de erro minimizada em RNAs, pode possuir mínimos locais e o método gradiente pode convergir a um deles.

Os resultados de uma RNA dependem da ordem de apresentação dos dados e dos pesos

iniciais atribuídos às suas conexões. Desta maneira, é necessário executá-la várias vezes com diferentes pesos e ordem dos dados para obter uma média do desempenho. As SVMs, não dependem da ordem com que os objetos são colocados para treinamento, porém, podem ter tempos maiores para predição do que RNAs (FACELI et al., 2011).

Portanto, as duas técnicas (RNA e SVM), cada qual com suas particularidades, apresentam bom desempenho preditivo.

2.6 AGRICULTURA DE PRECISÃO

A agricultura de precisão busca estudar e solucionar os problemas de variabilidade espacial, como variações de relevo, solo, vegetação e o histórico de uso das áreas de produção para gerenciar os sistemas agrícolas de forma localizada, com o objetivo de se obter resultados econômicos e ambientais sustentáveis. Para isto, utiliza-se de recursos de mapeamento das áreas, ferramentas de apoio a decisão e aplicação de recursos localizada (BERNARDI et al., 2014).

Ao utilizar as técnicas de AP, o produtor provavelmente terá melhores resultados econômicos, pois a partir delas é possível saber quais áreas devem ser priorizadas em cuidados agrônômicos, tratando os problemas de fertilidade do solo e da planta de forma específica e isolada para cada parte da área, retornando em menos desperdícios de insumos e maiores rendimentos da cultura (ANTUNIASSI et al., 2007).

A AP contribui para diminuição da degradação do meio ambiente. Isto se deve a racionalização de insumos agrícolas nas áreas de produção, pois se aplica os insumos somente nos locais que estão com deficiência ou escassez de nutrientes no solo e/ou na planta. Segundo Antuniassi et al. (2007), na Europa os maiores incentivos à prática de AP é em relação ao ganho ambiental e não econômico.

Recentemente, a AP tem se mostrado não só como um conjunto de ferramentas para manejo localizado da lavoura, como também uma nova forma de gestão e gerenciamento da produção agrícola, com um conjunto de tecnologias e procedimentos que otimizam o sistema de produção com o foco de estudar a variabilidade espacial da produção e seus fatores envolvidos.

Segundo Zhang et al. (2002), os fatores que influenciam a produção agrícola estão divididos em 6 grupos:

- Variabilidade da produção: distribuição histórica e atual da produtividade;

- Variabilidade do campo: altitude, declive, aspecto da paisagem, terraço, proximidade à bordadura e a rios e lagos;
- Variabilidade do solo: fertilidade do solo, atributos como de Nitrogênio (N), Fósforo (P), Potássio (K), Cálcio (Ca), Magnésio (Mg), Carbono (C), Ferro (Fe), Manganês (Mn), Zinco (Zn) e Cobre (Cu). Mudanças nos atributos de fertilidade do solo resultantes da aplicação de adubos orgânicos (ex. adubo verde, esterco), propriedades físicas do solo (textura, densidade, teor de umidade e condutividade elétrica), propriedades químicas (potencial de Hidrogeniônico (pH), Carbono orgânico, Capacidade de Troca de Cátions (CTC) e profundidade do solo);
- Variabilidade da cultura: densidade de plantio, altura da planta, estresse nutricional, estresse hídrico, propriedades biofísicas da planta (ex. índice de área foliar e biomassa), conteúdo de clorofila na folha e qualidade do grão;
- Variabilidade de fatores anômalos: infestação de plantas daninhas, ataque de pragas, presença de nematoides, geadas e granizo;
- Variabilidade no manejo: taxa de semeadura, rotação de culturas, aplicação de fertilizantes e pesticidas.

A Figura 27 mostra os fatores de variabilidade espacial de acordo com o grau de facilidade de intervenção.



Figura 27 – Exemplo de fatores envolvidos na variabilidade espacial de acordo com o grau de probabilidade de ocorrência e facilidade de correção.

Fonte: Molin (2003).

Devido à grande quantidade de variáveis que influenciam a produção agrícola, a AP surgiu como um método de administração da lavoura de forma cuidadosa e detalhada do solo e da cultura, que tem como objetivo adequar as diferentes condições de cada pedaço da área cultivada, levando em consideração a não uniformidade intrínseca do solo.

2.6.1 História da Agricultura de Precisão

A AP surgiu no final da década de 80, quando na Europa foi gerado o primeiro mapa de produtividade e nos EUA fez-se a primeira adubação com doses variadas, por meio da criação de uma ferramenta de aplicação localizada de fertilizantes que utilizava a amostragem das propriedades químicas do solo para aplicar as taxas variadas. Para aplicar os fertilizantes, utilizava-se bússola e cálculos de posição. Somente em 1990, surgiu o Sistema de Posicionamento Global (GPS), quando houve um considerável avanço nas pesquisas trazendo novos conhecimentos de informática e eletrônica aplicada (WHELAN; TAYLOR, 2013).

Em 1990 a AP chegou ao Brasil. Porém, a situação do Brasil era diferente da situação do local de origem da técnica. Seus recursos para investimento em mapas de colheita eram escassos e também havia falta de pessoas especializadas para o manuseio dos equipamentos de precisão. Ainda em 1995 era necessário importar equipamentos, especialmente as colhedoras equipadas com monitores de produtividade. Empresas privadas no Brasil e órgãos de pesquisa pública foram de extrema importância para que a técnica fosse difundida no Brasil (ACOSTA et al., 2010).

2.6.2 Ciclo da Agricultura de Precisão

Geralmente o ciclo começa pelo mapa de produtividade, obtido a partir de dados da colheita da cultivar, pois é necessário possuir dados do potencial produtivo da área cultivada, para avaliar, planejar e corrigir os problemas localizados.

O mapa de produtividade pode ser obtido de forma automática em colhedoras, mas para isto, é necessário alguns dispositivos especiais implantados na máquina. Porém, como alternativa, o mapa de produtividade pode ser gerado por meio de softwares, em que se insere os dados de cada ponto amostral e as coordenadas da área. Para obter os dados de cada ponto amostral sem o auxílio de colhedoras equipadas, realiza-se a colheita manual em cada ponto da lavoura, após a debulha manual, medição do teor de água e do peso dos grãos.

Molin (2001), MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (2011), afirmam que o mapa de produtividade é uma informação muito completa para se visualizar a variabilidade espacial das lavouras. Outras técnicas como fotografias aéreas, imagens de satélite, sistema de sensoriamento remoto, também são

utilizadas como ferramentas de diagnóstico da variabilidade espacial de uma lavoura.

A Figura 28 representa o ciclo, que se inicia na geração dos mapas de produtividade, após a amostragem do solo para análise dos componentes químicos, físicos e texturais. Com estes dados, realiza-se a análise dos resultados da produtividade relacionados às análises dos componentes do solo. Com isto, corrige-se o solo de acordo com as informações obtidas e o ciclo se reinicia.



Figura 28 – Ciclo da Agricultura de precisão.

Fonte: Molin (2001).

A análise dos solos é de extrema importância para saber se há quantidades de nutrientes disponíveis às plantas. Porém, há outros fatores que devem ser levados em conta, além da disponibilidade dos nutrientes às plantas, pois interferem na transferência do nutriente do solo até chegar a raiz da planta. Como exemplos, a quantidade de água disponível no solo, sua textura, as relações entre os nutrientes, a temperatura, a umidade do ar e do solo, a quantidade de energia solar disponível, entre outros (MOLIN, 2001).

Na próxima sessão serão detalhados os componentes do solo e sua influência na produtividade de uma cultura.

2.6.3 Componentes do solo e sua influência nas plantas

O solo influencia no crescimento das plantas devido aos fatores texturais do solo,

de composição mineralógica, de reação do solo (pH), do teor de matéria orgânica (MO), da disponibilidade de nutrientes, da precipitação, entre outros (NOVAIS et al., 2007; EPSTEIN; BLOOM, 2006).

A resistência a penetração do solo é um fator de influência na penetração das raízes, a qual depende da textura do solo (composição do solo em argila, areia e silte), da estrutura (densidade, porosidade, umidade, taxa de infiltração de água, erodibilidade, etc) e mineralogia (NOVAIS et al., 2007).

A água é o fator com maior influência no desenvolvimento da planta, pois a absorção do nutriente ocorre com o contato da água na raiz. Segundo Novais et al. (2007), os efeitos de reação do solo (pH) podem ser diretos como a disponibilidade dos elementos essenciais à nutrição da planta; solubilidade de elementos que podem ser tóxicos para as plantas; atividade de micro-organismos; favorecimento ou não de doenças nas plantas; habilidade de competição entre diferentes espécies de plantas e condições físicas do solo. Enquanto os efeitos indiretos, são relacionados a reação das propriedades químicas.

De acordo com a SOCIEDADE BRASILEIRA DE CIÊNCIA DO SOLO (2004) e Novais et al. (2007), para o desenvolvimento de culturas como a soja, os valores de pH adequados são entre 5.5 a 6.5.

A matéria orgânica é composta de resíduos de plantas e animais. Regula a disponibilidade de nutrientes, proporcionando melhora das propriedades físicas, químicas e biológicas, estabilizando as estruturas do solo e aumentando sua permeabilidade (COSTA et al., 2013).

Os macronutrientes são assim denominados, pois são os nutrientes necessários em quantidades relativamente grandes para as plantas, enquanto os micronutrientes, são necessários em quantidades comparativamente pequenas.

Dentre os vários macro e micronutrientes existentes, segundo Vitti e Trevisan (2000), a soja necessita do fornecimento dos seguintes nutrientes:

- N (fixação simbiótica, manejo da matéria orgânica),
- P, K, Ca, Mg, S, B, Cu, Mn, Zn, MO e Cobalto (Co) (fertilização mineral).

N e K são os nutrientes mais extraídos pela soja. P, Ca, Mg e S são os macronutrientes mais absorvidos pela planta.

O nitrogênio é necessário para síntese da clorofila e está envolvido no processo da fotossíntese. Sua deficiência, faz com que a planta deixe de absorver os nutrientes essenciais (NOVAIS et al., 2007). Segundo Epstein e Bloom (2006), a deficiência de N na planta é a que traz os piores efeitos, como crescimento retardado e aparência não viçosa.

O potássio participa do processo de fotossíntese da planta, contribui para o potencial

osmótico e está envolvido em muitos processos metabólicos, tendo grande impacto na produtividade da planta, afetando o incremento do peso e a qualidade de grãos de milho, e no caso da soja, afeta o conteúdo de óleo e proteínas. Também, auxilia na resistência a doenças na planta. Os sintomas de deficiência são necrose nas folhas mais velhas, retardo de crescimento e baixa resistência às doenças (NOVAIS et al., 2007; EPSTEIN; BLOOM, 2006).

Segundo Novais et al. (2007), o fósforo é muito importante na fotossíntese, respiração, armazenamento e transferência de energia, divisão e crescimento celular, entre outros processos que ocorrem nas plantas. Os sintomas de deficiência deixam as folhas torcidas com cor avermelhada ou púrpura, além de apresentar áreas mortas nas folhas nos frutos e no caule.

O cálcio é essencial para o crescimento de meristemas e para o crescimento e funcionamento dos ápices radiculares da planta. Auxilia na atividade microbiana, na disponibilidade de Molibdênio e na absorção de outros nutrientes. A deficiência de cálcio causa o apodrecimento apical e falência dos pontos de crescimento da planta (NOVAIS et al., 2007; EPSTEIN; BLOOM, 2006).

O magnésio difunde-se na suspensão celular. Sua propriedade mais importante é a solubilidade de seus compostos. Participa de reações enzimáticas como na transferência de fosfato e nucleotídeos. Influencia no movimento de carboidrato das folhas para outras partes da planta e estimula a captação de transporte de fósforo na planta. Sua deficiência ocorre na forma de clorose entre as nervuras, na diminuição da acidez total e vitamina C nas plantas e redução dos frutos (NOVAIS et al., 2007).

O último macronutriente mais absorvido pela soja a ser comentado é o enxofre. É um elemento importante para a produção de aminoácidos, proteínas e clorofila. Melhora o crescimento de raízes e conseqüentemente o vigor da planta (NOVAIS et al., 2007). Segundo Epstein e Bloom (2006), os sintomas de deficiência do enxofre são muito semelhantes aos de nitrogênio, pois ambos são constituintes de proteínas, porém, a clorose causada pela deficiência de enxofre, ocorre nas folhas jovens e não nas maduras como ocorre na deficiência de nitrogênio.

A Tabela 2 mostra as quantidades de nutrientes absorvidas e exportadas nos grãos pela cultura da soja.

Segundo Vitti e Trevisan (2000), na época de pré-plantio, Mg pode ser fornecido ao solo na forma de calagem, e Ca na forma de calagem e gessagem, assim como enxofre pode ser colocado no solo por meio de gessagem e o P e K na forma de fosfatagem e potassagem, respectivamente.

Tabela 2 – Quantidade absorvida e concentração de nutrientes pela cultura da soja.

Parte da Planta	N	P	K	Ca	Mg	S	B	Cl	Mo	Fe	Mn	Zn	Cu
	kg (1000 kg ⁻¹) ou kg ⁻¹						g (1000 kg ⁻¹) ou g kg ⁻¹						
Grãos	51	5	17	3	2	5,4	20	237	5	70	30	40	10
Restos Culturais¹	31	2,5	7,5	9,2	4,7	10	57	278	2	390	100	21	16

Folhas pecíolos e caules que são restituídos ao solo.

Fonte: Adaptado de Vitti e Trevisan (2000), Gianluppi et al. (2009).

A diagnose foliar, segundo Gianluppi et al. (2009), é uma ferramenta de interpretação dos dados da análise do solo, pois analisa quimicamente as folhas e os resultados podem ser interpretados conforma a Tabela 3.

Tabela 3 – Concentrações de nutrientes utilizadas para interpretação dos resultados de análise foliar da soja no início do florescimento.

Elementos	Deficiente ou		Suficiente ou		Excessivo ou Muito Alto
	Muito baixo	Baixo	Médio	Alto	
	g kg ⁻¹				
N	<32,5	32,5 - 45,0	45,1 - 55,0	55,1 - 70,0	>70,0
P	<1,6	1,6 - 2,5	2,6 - 5,0	5,1 - 8,0	>8,0
K	<12,5	12,5 - 17,0	17,1 - 25,0	25,1 - 27,5	>27,5
Ca	<2,0	2,0 - 3,5	3,6 - 20,0	20,1 - 30,0	>30,0
Mg	<1,0	1,0 - 2,5	2,6 - 10,0	10,1 - 15,0	>15,0
S	<1,5	1,5 - 2,0	2,1 - 4,0	>4,0	-
	mg kg ⁻¹				
Mn	<15,0	15,0 - 20,0	21,0 - 100,0	101,0 - 250,0	>250,0
Fé	<30,0	30,0 - 50,0	51,0 - 350,0	352,0 - 500,0	>500,0
B	<10,0	10,0 - 20,0	21,0 - 55,0	56,0 - 80,0	>80,0
Cu	<5,0	5,0 - 9,0	10,0 - 30,0	31,0 - 50,0	>50,0
Zn	<11,0	11,0 - 20,0	21,0 - 50,0	51,0 - 75,0	>70,0
Mo	<0,5	0,5 - 0,9	1,0 - 5,0	5,1 - 10,0	>10,0

Fonte: Adaptado de Gianluppi et al. (2009).

Caso haja a deficiência de algum nutriente, segundo os autores Gianluppi et al. (2009), dificilmente será possível a correção na mesma safra, porém pode efetuar a correção de calcário e adubos na próxima safra.

Para realizar a análise foliar da soja, deve ser coletada a terceira ou quarta folha a partir do ápice da planta, e no mínimo 40 folhas por talhão no início da floração. Antes de serem levadas ao laboratório de análises foliares, deve-se deixar secar as folhas à sombra e em sacos de papel (GIANLUPPI et al., 2009).

2.6.4 Aplicação da Inteligência Artificial na agricultura

O trabalho de Soares et al. (2014), utilizou redes neurais artificiais para predição da produção na cultura do feijoeiro, com o objetivo de avaliar o desempenho de redes neurais para predição da produtividade do feijão. Para treinamento e validação, utilizou-se dados morfológicos da cultura.

Testou-se várias arquiteturas de redes neurais do tipo perceptron multicamadas com o algoritmo *Backpropagation*, variando o número de neurônios da camada escondida. As variáveis de entrada da rede neural eram: matéria seca total (g); altura de planta (cm) e n° de plantas por m^2 . Na camada de saída, utilizou-se a variável de resultado da produção de grãos ($kg\ ha^{-1}$) correspondente as entradas.

Escolheu-se das arquiteturas treinadas, aquela com menor erro relativo médio e menor variância em comparação aos dados de validação. A rede que obteve o melhor desempenho no estudo, foi a rede com 30 neurônios na camada escondida (Figura 29).

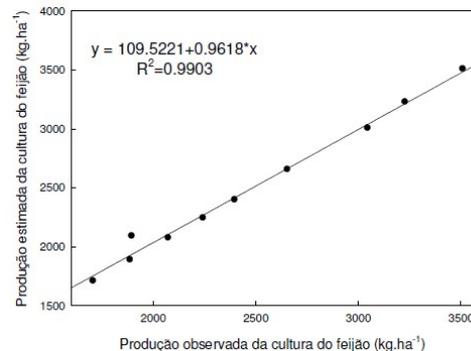


Figura 29 – Comparação dos valores reais e estimados da cultura do feijão pela melhor rede neural encontrada.

Fonte: Soares et al. (2014).

O trabalho dos autores Leal et al. (2015), tinha como meta encontrar entre redes neurais e a técnica de regressão múltipla, a que melhor estimasse a produtividade do milho, utilizando como dados de treinamento e validação, os atributos do solo: MO, CTC, saturação de bases (V%), e teor de argila (TA).

Utilizou-se o método dos mínimos quadrados ordinários, para estimar a produtividade do milho via regressão linear múltipla. Para utilizar a rede neural artificial no trabalho, escolheram o algoritmo de treinamento *resilient propagation*, para RNAs do tipo *Multilayer Perceptron*. A rede neural possuía uma camada escondida, composta por 4 neurônios, sendo assim, a RNA ficou com 4 neurônios na camada de entrada e escondida e um na camada de

saída. A função de ativação utilizada foi a sigmoïdal.

Em comparação às RNAs, a técnica de regressão múltipla apresentou inferioridade nas estatísticas utilizadas para medir o desempenho de previsão da produtividade do milho e tendia a superestimar as menores classes e subestimar as maiores classes de produtividade.

Bocca (2014), realizou o estudo: Produtividade de cana-de-açúcar: caracterização dos contextos de decisão e utilização de técnicas de mineração de dados para modelagem. O objetivo do seu trabalho era modelar numericamente a produtividade da cana de açúcar, considerando dados relativos a solo, manejo e clima em um contexto do cultivo não instalado.

As técnicas utilizadas foram redes neurais, árvore de modelos, regressão múltipla, *Stepwise*, árvore de regressão, SVM e *Random Forest*. Os melhores desempenhos de modelagem de previsão foram das técnicas de SVM e *Random Forest*. No entanto, a técnica de *Random Forest* utilizou quase o dobro de variáveis em relação à SVM para a previsão da produtividade.

Rub (2009) realizou um trabalho para obter uma técnica de previsão da produtividade do trigo. Para isto utilizou sete variáveis de entrada, sendo aplicações de nitrogênio em três pontos diferentes da área cultivada, reflexão espectral da planta em duas fases de crescimento do trigo - *Red Edge Inflection Point 32* e *49 (REIP)*, além da condutividade elétrica do solo.

Utilizou o modelo de rede neural MLP e RBF, Árvore de regressão *top-down* e SVM. Para a rede neural MLP, utilizou-se 10 neurônios na camada escondida e foi treinada até obter um erro de 0,001. Para a rede de função de base radial, um raio de 1 foi utilizado para a base radial de neurônios na camada oculta em que o algoritmo acrescentava neurônios até no máximo 70, buscando obter o erro de 0,001. Na configuração da árvore de regressão, foram utilizadas os ajustes padrão. Na SVM, o melhor resultado se obteve com a utilização do kernel de função de base radial com *trade-off* de 60, desvio padrão 4, para a variável de folga 0,2. Como resultado, a SVM foi a técnica que apresentou melhor desempenho e também o menor tempo de computação. Já a árvore de regressão apresentou o pior desempenho de todas as técnicas utilizadas.

3 MATERIAIS E MÉTODOS

Após estudos realizados sobre análise foliar e aplicações de técnicas de inteligência artificial na agricultura de precisão, é importante definir as ferramentas e procedimentos a serem utilizados para o desenvolvimento da pesquisa. Portanto, neste capítulo serão descritos os materiais e métodos utilizados para alcançar o objetivo deste trabalho.

3.1 LOCALIZAÇÃO DO EXPERIMENTO

Para obtenção dos dados utilizados neste trabalho foram realizadas 3 coletas das folhas de soja em duas áreas experimentais (Talhão 1, com 10 ha e Talhão 2, com 23,8 ha, Figura 30), ambas cultivadas sob sistema de plantio direto por pelo menos 10 anos com sucessão de cultura soja/milho para fins comerciais.

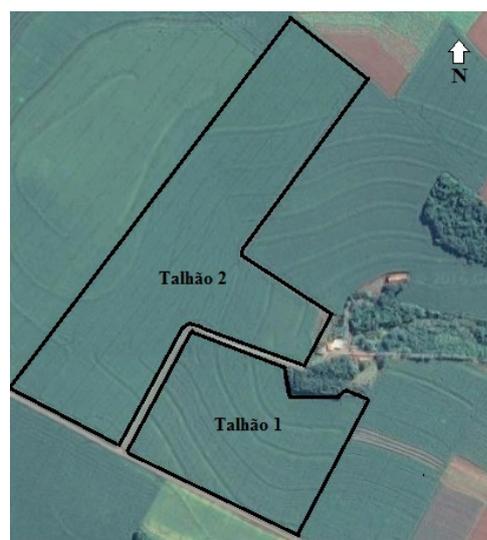


Figura 30 – Áreas experimentais.

Fonte: Autoria própria

As áreas estão localizadas no município de Serranópolis do Iguçu/PR, sob coordenadas geográficas 25°24'28" S e 54°00'17" O com elevação média de 355 m. Foram determinados por meio do software Pathfinder¹ versão 4.0, duas grades amostrais irregulares, contendo 42 pontos amostrais no talhão 1 e 73 no talhão 2 (Figura 31), com densidade de aproximadamente 3 pontos por ha.

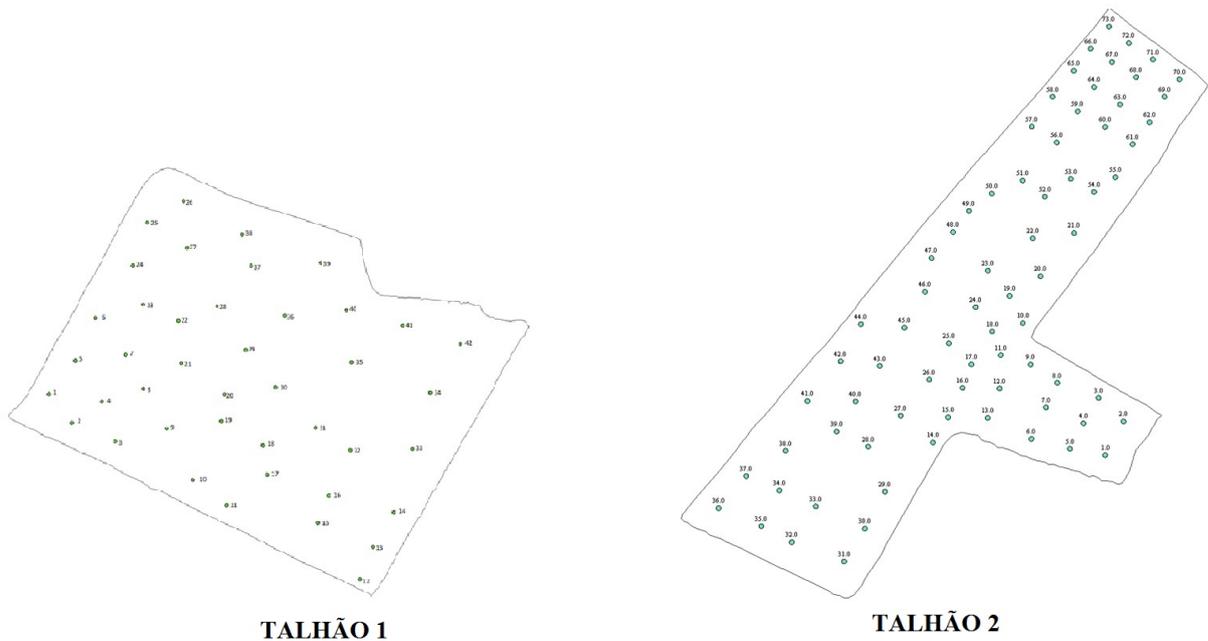


Figura 31 – Grades de amostragens dos talhões.

Fonte: Autoria própria

3.2 MATERIAIS PARA COLETA DA FOLHA DA SOJA E PARA ANÁLISE FOLIAR

A coleta dos dados foi realizada de forma manual no ano de 2012, em 3 estádios da safra da soja 2012/2013, V3, V6 E R1, onde o estádio V3 é o estádio vegetativo da planta com 3 folhas, o V6 é o estádio da planta com 6 folhas e o estádio reprodutivo da planta já com flores, chamado de R1 (WOLLENHAUPT et al., 1997). Os pontos foram localizados por meio de um receptor GPS (GPS Trimble Pathfinder Pro XT). Para cada ponto amostral em um raio de 3 metros, coletou-se 12 folhas (WOLLENHAUPT et al., 1997). As folhas foram acondicionadas

¹<http://www.trimble.com/mappingGIS/PathfinderOffice.aspx>

em sacolas de papel para posteriormente serem secadas em estufa a 105°. Após a secagem, as folhas foram debulhadas por um moinho de facas modelo SL 31 da marca SOLAB (Figura 32). Ao final de todo este processo as folhas foram encaminhadas para análise laboratorial dos macronutrientes Ca, Mg, P, K, N, no Laboratório de Solos da Universidade Tecnológica Federal do Paraná (UTFPR) - Câmpus Pato Branco em convênio com o Instituto Agronômico do Paraná (IAPAR), os quais seguem a metodologia de Tedesco et al. (1995) para a determinação de macronutrientes em tecido vegetal.



Figura 32 – SL-31 - Moinho de facas tipo Willey.

Fonte: SOLAB - Equipamentos para laboratório (2016)

Para determinar a produtividade da soja 2012/2013 também foi realizada coleta de forma manual, nos meses de janeiro/fevereiro, coletando-se para cada ponto amostral, duas linhas de plantio, por um percurso de 1 m, representando uma porção de cerca de 0,9 m². Após a coleta, foi realizada a debulha manual e realizou-se o procedimento de secagem em estufa a 105° para posteriormente ser calculada a produtividade em toneladas por hectare (ton ha⁻¹), considerando uma umidade de grãos de 12%.

3.3 TECNOLOGIAS UTILIZADAS

O ambiente de desenvolvimento utilizado para executar os testes foram duas máquinas virtuais criadas em um dos servidores da UTFPR (disponibilizado pelo professor orientador

deste trabalho) com o sistema operacional Microsoft Windows 7, as quais operam sobre a plataforma de virtualização Citrix XenServer² versão 6.2.

Esta plataforma foi desenvolvida em 2003 como um projeto de pesquisa na Universidade de Cambridge, liderado por Ian Pratt e Simon Crosby, professores na Universidade de Cambridge. A plataforma tem sido mantida pela empresa norte-americana Citrix Systems (CITRIX SYSTEMS INCORPORATION, 2015).

O XenServer é um software de código aberto utilizado para segurança de vários sistemas operacionais e aplicativos em um único dispositivo, reduzindo custos e simplificando o gerenciamento da tecnologia de informação de servidores e estruturas virtuais em nuvem e desktop.

O tipo de rede neural construída neste trabalho é do tipo perceptron multicamadas com o algoritmo de treinamento *Backpropagation* e para implementar a RNA utilizou-se o software Stuttgart Neural Network Simulator³ (SNNS), desenvolvido na Universidade de Stuttgart na Alemanha em 1989 com o objetivo de criar um ambiente de simulação eficiente e flexível para a pesquisa e aplicação de redes neurais.

O SNNS possui um simulador de kernel escrito em C, uma interface gráfica que pode ser usada para criar diretamente, manipular e visualizar redes neurais em várias formas. Também possui um interpretador de linha de comando e uma linguagem de programação chamados batchman e um compilador de redes neurais – snns2c.

Através do SNNS, pode se definir funções de ativação, funções de inicialização, funções de aprendizagem, funções de atualização, implementadas como programas em C simples e ligadas ao kernel do simulador (MAMIER; WIELAND, 1995).

Outra aplicação utilizada foi o Java Neural Network Simulator⁴ (JavaNNS), apenas para visualizar a rede treinada no SNNS, pois não possui formas de implementar redes neurais de forma a automatizar a criação, testes e validação de RNAs. O JNNS foi desenvolvido pela Universidade de Tübingen da Alemanha por volta dos anos de 1995 a 2000. O JavaNNS é um software sucessor do SNNS que possui uma interface gráfica desenvolvida na linguagem Java, sendo compatível com todas as funcionalidades do SNNS. A Figura 33 mostra a interface gráfica de visualização do JavaNNS.

As SVMs foram criadas utilizando a funções kernel polinomial. Para desenvolver as SVMs, utilizou-se a Application Programming Interface⁵ (API) da Waikato Environment for Knowledge Analysis⁶ (WEKA), desenvolvida inicialmente em 1992 pela universidade de Waikato da Nova Zelândia (HALL et al., 2009). Por meio da WEKA, se tem acesso a testes

²<https://www.citrix.com.br/>

³<http://www.ra.cs.uni-tuebingen.de/SNNS/>

⁴<http://www.ra.cs.uni-tuebingen.de/software/JavaNNS/>

estatísticos do aprendizado de máquina incluídos na versão 3.6.

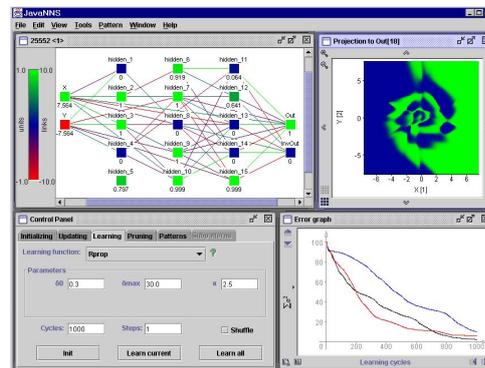


Figura 33 – Interface do software JavaNNS.

Fonte: Tübingen (2016)

A WEKA também foi utilizada como ferramenta para seleção de variáveis. Desta forma, realizou-se a seleção de atributos para verificar se o desempenho do aprendizado poderia diminuir ou continuar satisfatório com menos atributos de entrada, para conseqüentemente baratear o custo para futuros aplicativos destas técnicas de IA.

A escolha da ferramenta WEKA deveu-se ao seu conjunto de algoritmos de aprendizado de máquina para tarefas de mineração de dados, os quais podem ser aplicados diretamente a um conjunto de dados ou chamado a partir de seu próprio código Java. Weka contém ferramentas para pré-processamento de dados, classificação, regressão, clustering, regras de associação, e visualização, as quais são úteis para o desenvolvimento deste trabalho.

Para determinação dos pontos amostrais das áreas experimentais, utilizou-se o software Pathfinder, pois é um software fácil de utilizar e possui ferramentas de pós-processamento *Global Navigation Satellite Systems* (GNSS). É um software projetado para desenvolver informações Geographic Information System (GIS) que são informações consistentes, confiáveis e precisas de dados GNSS coletados em campo. O software faz correções diferenciadas para melhorar a qualidade dos dados GNSS coletados em campo, realiza processamento de dados em alta precisão e é capaz de importar e exportar dados de uma variedade de formatos de GIS.

Com o GPS foi realizado o percurso do contorno da área e utilizou-se o programa auxiliar da ferramenta para fazer a divisão da área em pontos amostrais, com uma média aproximada de pontos por hectares estabelecidos.

⁵<https://weka.wikispaces.com/Use+WEKA+in+your+Java+code>

⁶<http://www.cs.waikato.ac.nz/ml/weka/>

3.4 PRÉ-PROCESSAMENTO DOS DADOS

Para aplicar as técnicas de Inteligência Artificial foi necessário realizar a normalização dos dados e a modelagem de acordo com o formato requerido por cada ferramenta a ser utilizada.

3.4.1 Normalização dos dados

Devido a grande amplitude dos dados de entrada e saída utilizados nos testes, foi necessário fazer um tratamento dos dados para que estes ficassem em uma pequena amplitude antes de processá-los em uma RNA ou SVM. Como os dados utilizados neste trabalho são dados numéricos e a função de ativação (no caso das RNAs) utilizada é a de sigmoide logística (função em que seus valores são limitados pelo intervalo [0 e 1]). Deste modo, valores muito altos e muito baixos em um mesmo conjunto de dados podem saturar a função de ativação e tornar o aprendizado mais lento, portanto, os dados foram normalizados no intervalo de 0 a 1 de acordo com a Equação 18 antes de utilizá-los nas ferramentas computacionais (SU et al., 1997).

$$f(x) = y = (m_{max} - m_{min}) \cdot \frac{(l - l_{min})}{(l_{max} - l_{min})} + m_{min} \quad (18)$$

Onde m é o valor normalizado correspondente ao valor original l , e l_{min} e l_{max} são os valores de mínimo e de máximo entre todos os valores da base de dados e m_{min} corresponde a 0 e m_{max} corresponde a 1.

Após a normalização os dados foram formatados de acordo com cada ferramenta utilizada.

A escolha pela função sigmoide logística se deve por ser a função de ativação mais utilizada e por oferecer a importante flexibilidade dada por sua inclinação na origem. Além disso, a função logística apresenta propriedades de simetria e de completa diferenciabilidade (XAVIER, 2005).

3.4.2 Modelagem dos dados

Para utilizar o software SNNS e JNNS, os dados foram formatados como mostra a Figura 34, em que deve haver no começo do arquivo a especificação do número de amostras; o número de parâmetros de entrada; e número de parâmetros de saída. Estes arquivos devem ser nomeados com extensão pat. As entradas devem ser separadas por um espaço em uma linha e o(s) parâmetro(s) de saída na linha seguinte à(s) suas respectivas entradas. O comentário é representado pelo símbolo # a sua frente.

Os dados utilizados na ferramenta WEKA e IDE do NetBeans (utilizando API da WEKA) para criação dos modelos de SVMs e seleção de atributos foram formatados como arquivos ARFF padronizados conforme a Figura 35, onde o @ indica início de um comentário. As vírgulas separam as variáveis, sendo a última variável da linha, a variável de saída.

```
SNNS pattern definition file V3.2
generated at Mon Apr 25 17:54:03 1994

No. of patterns : 70
No. of input units : 5
No. of output units : 1

#Amostra A27 C2
0.7722 0.8710 0.6675 0.3324 0.3662
0.4535
#Amostra B26 C2
0.7921 0.1722 0.6038 0.5680 0.4465
0.2111
#Amostra B42 C2
0.0162 0.2437 0.4047 0.7253 0.6168
0.4915
#Amostra B47 C2
0.7921 0.2594 0.6038 0.4984 0.3846
0.5743
#Amostra B9 C2
0.8970 0.1633 0.4047 0.1575 0.6871
0.6134
```

Figura 34 – Arquivo de entrada no software SNNS e JavaNNS para treinamento de uma rede neural.

Fonte: Autoria própria.

```
@relation produtividade
@attribute N numeric
@attribute P numeric
@attribute K numeric
@attribute Ca numeric
@attribute Mg numeric
@attribute Prod numeric

@data
0.8306,0.4451,0.5758,0.4114,0.3787,0.2111
0.7879,0.5372,0.7188,0.6596,0.6287,0.8368
0.8306,0.3194,0.5758,0.4477,0.5496,0.5394
0.7859,0.6768,0.5758,0.0018,1.0285,0.5422
0.9151,0.4354,0.5758,0.502,0.4267,0.4971
0.8732,0.5233,0.8617,0.5922,0.4326,0.1684
0.0116,0.4646,0.4329,0.3367,0.7453,0.5765
0.9159,0.3724,0.4329,0.4013,0.3912,0.1564
0.829,0.7734,0.5758,1.0013,0.9751,0.5384
0.8732,0.5177,0.5758,0.558,0.5496,0.0463
0.7879,0.2858,0.2899,0.6103,0.6485,0.3404
```

Figura 35 – Formato de um arquivo de entrada no software WEKA.

Fonte: Autoria própria.

3.5 CONJUNTO DE DADOS PARA TREINAMENTO E VALIDAÇÃO

Os dados utilizados como entrada eram dos nutrientes N, P, K, Ca e Mg e o valor de saída era o da produtividade da soja. Desta maneira, o conjunto de treinamento representava uma porção de 60% do conjunto total, enquanto o conjunto de validação representou 30% dos dados e o conjunto de teste, os outros 10% restantes, pois conforme SILVA et al. (2010), uma estratégia muito comum para a realização do particionamento desses subconjuntos é dividir os dados em um conjunto maior, frequentemente 60 a 90%, para treinar a rede neural e estimar seus parâmetros e o conjunto de dados restante, e menor, para testar o modelo.

Dentre os subconjuntos divididos, o subconjunto de validação foi utilizado dentro da rede e da SVM logo após o treinamento, para verificar se a técnica aprendeu como obter a saída por meio das entradas. O subconjunto de teste, foi utilizado após escolher a rede com menor erro para somente aplicar as entradas deste conjunto no algoritmo e após verificar se as saídas estimadas foram significativamente semelhantes ou não.

O erro analisado nas RNAs e SVMs foi o *Mean squared error* (MSE) (Equação 19).

$$MSE = \frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (19)$$

Onde N é o número total de amostras, \hat{y}_i é o valor simulado e y_i o valor real.

3.6 ANÁLISE E COMPARAÇÃO DAS TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL

Para realizar a escolha de uma RNA e uma SVM que atendessem ao propósito deste trabalho, foi realizada a validação cruzada e análises estatísticas.

3.6.1 Critério de parada do treinamento

A técnica utilizada para determinar o momento de parar o treinamento das técnicas de inteligência artificial foi a do treinamento com parada antecipada (HAYKIN, 2001), a qual permite comparar a cada ciclo de treinamento o valor estimado pela RNA ou SVM com o valor real da produtividade da soja. A partir disto, obtêm-se um erro, se o erro for tão baixo quanto aceitável, a RNA ou SVM treinada foi satisfatória e encerram-se os treinamentos, caso contrário, deve ser continuado o treinamento até que se obtenha o erro esperado, ou até que o erro de validação comece a aumentar.

3.6.2 Análise Estatística

Para avaliar o modelo baseado nas técnicas de inteligência artificial resultante deste trabalho utilizou-se: teste de normalidade dos dados para confirmar se os testes estatísticos paramétricos eram aplicáveis neste estudo; medida da variabilidade dos dados como desvio padrão e variância; medida de tendência central; teste de hipóteses H_0 e H_1 para testar se os dados reais e os simulados poderiam ser considerados provenientes de uma mesma população (o que é esperado); coeficiente de determinação R^2 para testar o quanto o modelo simulado se ajustou aos dados reais e o coeficiente de correlação de Pearson $R\rho$ para quantificar a correlação dos dados reais e simulados.

4 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados os resultados obtidos das RNAs e SVMs implementadas, uma discussão individual de seus desempenhos e a comparação das duas técnicas bem como a viabilidade ou não de se utilizar os modelos criados para predição da produtividade da soja somente com os nutrientes da folha.

4.1 ANÁLISE EXPLORATÓRIA DOS DADOS

Os dados foram analisados previamente antes de aplicar testes estatísticos para entendimento básico dos dados e das relações existentes entre eles. As Tabelas 4, 5, 6, 7 e 8 mostram a estatística descritiva dos dados para os nutrientes Nitrogênio, Potássio, Fósforo, Magnésio e Cálcio, respectivamente.

Tabela 4 – Estatística descritiva do Nitrogênio separada por coletas e áreas.

Atributo	Área	Coleta	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
N%	A	C1	3,84	0,68	17,75	0,67	3,84	4,83	3,57	10,58
		C2	3,83	1,94	50,63	0,43	4,65	5,9	4,65	-0,53
		C3	4,04	1,12	27,72	0,39	4,29	5,36	4,29	6,3
	B	C1	3,68	1,12	30,64	0,64	4,02	4,82	4,11	3,16
		C2	3,65	1,68	46,02	0,42	4,47	5,54	4,47	-0,09
		C3	4,19	0,55	13,19	1,07	4,29	5,18	4,11	14,29

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

Analisando a média de nitrogênio (Tabela 4), em ambas as áreas a terceira coleta teve médias maiores que as demais coletas. Porém, na análise da mediana, a coleta 2 teve maiores quantidades de nitrogênio em ambas as áreas, assim como os valores de moda. O desvio padrão em torno da média é de até 1,94%. O Coeficiente de Variação (CV) ficou classificado como

muito alto em ambas as áreas segundo Pimentel e Garcia (2002). Os valores de máximo também se concentraram na coleta 2 nas duas áreas amostrais. Para a coleta 2 a curtose se apresentou como leptocúrtica e para as demais como platicúrtica.

A média do potássio (Tabela 5) varia entre 1 e 2 nas duas áreas amostrais, com um desvio padrão de até 0,4. O CV se apresentou de médio a alto. A curtose se apresentou como leptocúrtica na coleta 3 da área A e coleta 2 da área B e platicúrtica nas demais.

Tabela 5 – Estatística descritiva do Potássio separada por coletas e áreas.

Atributo	Área	Coleta	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
K%	A	C1	2,26	0,33	14,72	1,18	2,29	3,13	2,29	2,19
		C2	1,60	0,4	24,47	0,06	1,74	2,57	1,74	5,28
		C3	1,13	0,24	21,71	0,62	1,18	1,74	1,18	-0,15
	B	C1	2,13	0,39	18,21	1,18	2,01	3,13	2,29	0,25
		C2	1,59	0,34	21,39	0,90	1,74	2,29	1,74	-0,44
		C3	1,26	0,32	25,66	0,06	1,18	2,01	1,18	2,41

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

A média de fósforo (Tabela 6) para as três coletas é em quantidades menores que 1%. O desvio padrão é um pouco grande em torno da média de fósforo na terceira coleta da área B. O CV varia de médio a alto. Os valores de máximo não passaram de 1. A distribuição se apresentou como platicúrtica para as coletas das áreas A e B, com exceção das coletas 1 e 2 da área A, que se apresentaram com curtose leptocúrtica.

Tabela 6 – Estatística descritiva do Fósforo separada por coletas e áreas.

Atributo	Área	Coleta	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
P%	A	C1	0,65	0,12	19,13	0,38	0,63	0,97	0,75	0,11
		C2	0,47	0,05	10,91	0,35	4,65	0,55	0,45	-0,06
		C3	0,39	0,04	10,82	0,26	0,40	0,46	0,41	1,41
	B	C1	0,65	0,08	12,87	0,40	0,64	1,00	0,64	4,20
		C2	0,48	0,1	20,32	0,33	0,47	1,08	0,46	19,63
		C3	0,43	0,69	16,14	0,21	0,42	0,70	0,48	6,71

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

A média de magnésio (Tabela 7) está em torno de 0,3 com um pequeno desvio padrão variando de 0,3 a 0,5. O CV varia de médio a alta variação dos dados. Os valores de máximo se aproximam dos valores de média, assim como a moda. A distribuição se apresenta platicúrtica

em todos os casos com exceção da área A na primeira coleta, que apresenta distribuição leptocúrtica.

Tabela 7 – Estatística descritiva do Magnésio separada por coletas e áreas.

Atributo	Área	Coleta	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
Mg%	A	C1	0,33	0,05	15,03	0,21	0,33	0,42	0,31	-0,48
		C2	0,26	0,03	11,06	0,21	0,26	0,36	0,26	1,70
		C3	0,20	0,03	17,32	0,11	0,20	0,30	0,21	1,06
	B	C1	0,33	0,04	12,65	0,19	0,33	0,44	0,31	1,69
		C2	0,26	0,04	16,35	0,14	0,26	0,37	0,23	0,78
		C3	0,22	0,05	22,88	0,04	0,21	0,36	0,18	2,19

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

A média de cálcio (Tabela 8) na folha fica em torno de 0,7 a 1,14%, o desvio padrão varia de 0,27 a 0,4 em torno da média, enquanto o CV apresenta alta a muito alta a variação dos dados. A mediana apresenta valores mais altos que a média e os valores de moda. A distribuição classifica-se como platicúrtica nas coletas 1, 2 e 3 da área A e coleta 1 da área B. Nas coletas 1 e 2 da área B classifica-se como uma distribuição leptocúrtica.

Tabela 8 – Estatística descritiva do Cálcio separada por coletas e áreas.

Atributo	Área	Coleta	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
Ca%	A	C1	0,74	0,37	50,50	0,02	0,85	1,95	0,81	1,67
		C2	0,77	0,3	38,55	0,20	0,83	1,74	0,69	2,06
		C3	1,23	0,33	26,95	0,29	1,31	1,71	1,42	2,36
	B	C1	0,96	0,27	28,55	0,36	0,98	1,67	0,98	0,31
		C2	0,64	0,3	46,64	0,08	0,73	1,15	0,46	-0,87
		C3	1,14	0,4	35,29	0,28	1,26	1,75	0,29	-0,16

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

A média de produtividade da área A ficou em torno de 5,07 ton/ha, enquanto a produtividade da soja na área B foi inferior a área A com 4,81 ton/ha, o que também ocorreu com a mediana. O desvio padrão das duas áreas esteve igual, variando uma casa decimal, apenas. O CV das duas áreas também teve uma variação muito semelhante de classificação média. Apesar de a maior média de produtividade ser na área A, o menor valor de produtividade também se encontrou na área A, e o valor maior de produtividade foi encontrado na área B. Na área A não teve valor de produtividade que se repetiu, e para a área B, o valor de moda foi pouco menor que o seu valor de média. A distribuição da área A classificou-se como

platicúrtica e a distribuição da área B como leptocúrtica (Tabela 9).

Tabela 9 – Estatística descritiva da produtividade da soja separada por áreas.

Atributo	Área	\bar{x}	S	CV	Min.	Me	Max.	Mo	K
Produtividade (ton/ha)	A	5,07	0,7	13,75	2,42	5,06	6,61	*	4,28
	B	4,81	0,69	14,29	3,63	4,75	6,69	4,34	-0,27

\bar{x} - Média; S - Desvio Padrão; CV - Coeficiente de Variação; Min. - Mínimo; Max. - Máximo; Mo - Moda; K - Curtose.

Fonte: Autoria própria

A construção dos boxplots realizados para cada coleta em ambas as áreas, mostrou que todos os atributos analisados e a produtividade possuíam algum ponto discrepante, com exceção da terceira coleta do potássio, que não apresentou pontos discrepantes.

Realizou-se os testes de normalidade dos dados dos nutrientes e das produtividades pelos testes de Anderson - Darling, Kolmogorov - Smirnov, Shapiro - Wilk, Ryan - Joiner. Para os atributos, nitrogênio, potássio e cálcio, não foi possível verificar normalidade ao nível de 5% ($\alpha = 0,05$) de significância. Portanto, buscou-se transformar os dados pela transformação logarítmica, raiz quadrada dos dados e a divisão de $1/x$ chamada de transformação do quociente. Porém, como resultado, os testes de normalidade apresentaram valores de p mais distantes do nível de 5% de significância do que os valores de p obtidos antes da transformação dos dados.

Portanto, rejeita-se a hipótese de que os dados dos nutrientes seguem uma distribuição normal. A não normalidade dos dados é devido a deficiência ou excesso de nutrientes ou de elementos químicos nocivos à planta. Deste modo, considerando que a normalidade dos dados não pode ser assumida, os testes de avaliação das técnicas de inteligência artificial a serem apresentados foram baseados em métodos estatísticos não paramétricos.

4.2 DESEMPENHO DAS REDES NEURAIS ARTIFICIAIS

Criou-se Redes Neurais Artificiais do tipo Perceptron Multicamadas utilizando o algoritmo *Backpropagation* para a primeira coleta dos nutrientes da folha da soja no estádio V3, para a segunda coleta no estádio V6 e para a terceira coleta - estádio R1. Utilizou-se topologias de redes com até três camadas intermediárias, variando a quantidade de neurônios de

1 a 500 em cada camada, totalizando aproximadamente 12000 redes neurais artificiais treinadas. A primeira camada da rede neural, recebeu como entradas os nutrientes N, P, K, Ca e Mg em porcentagem e a última camada, a camada de saída, o valor da produtividade da soja em toneladas por hectare, porém, estes valores foram todos normalizados entre 0 e 1 antes de serem utilizados na RNA.

Os resultados apresentados a seguir, são das redes que tiveram o menor MSE. Conforme a Tabela 10, verifica-se quais foram as topologias das redes neurais com menor MSE, a quantidade de épocas até o momento de parada da execução da RNA (quando o erro de validação começa a subir), os valores de MSE, os coeficientes de determinação (R^2) e os coeficientes de correlação de Pearson (R_p) separados para cada coleta.

Tabela 10 – Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (pontos amostrais aleatórios para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Topologia	Épocas	MSE	R^2	R_p
1	1x5_1x1_1x1	67	0,0387	0,7407	-0,8606
2	1x5_1x6_1x1	15	0,0242	0,0085	0,0922
3	1x5_1x5_1x1	60	0,0141	4E-05	-0,0063

O número que antecede a X representa o número de camadas e o valor posterior a X o número de neurônios em cada camada, onde o primeiro trio de caracteres representa a camada de entrada, o segundo trio de caracteres a camada intermediária e o último a camada de saída.

Fonte: Autoria própria.

Apesar de ter sido realizados testes com até três camadas intermediárias, os melhores desempenhos foram com uma camada intermediária e com uma pequena quantidade de neurônios, variando de um neurônio somente na RNA da primeira coleta, 6 neurônios na camada intermediária da RNA da segunda coleta e 5 neurônios na camada intermediária da RNA da terceira coleta. Como provou Cybenko (1989), a partir de extensões do Teorema de Kolmogoroff, são necessárias no máximo duas camadas intermediárias, com um número suficiente de unidades por camada, para se produzir quaisquer mapeamentos. Também foi provado que apenas uma camada intermediária é suficiente para aproximar qualquer função contínua, deste modo, os testes realizados foram suficientes para verificar a capacidade de aprendizado de uma rede neural com os dados utilizados neste trabalho.

O coeficiente de determinação é utilizado para avaliar a qualidade do ajuste do modelo, pois indica quanto o modelo foi capaz de explicar os dados coletados. Portanto, quanto mais próximo de 1, melhor o modelo é. A primeira coleta dos nutrientes da folha foi a única que teve um bom desempenho, pois a melhor RNA obtida com os dados da primeira coleta teve um modelo que conseguiu explicar 74% da variação dos dados reais. As demais coletas tiveram

modelos muito ruins, que não conseguiram explicar nem 1% da variação dos dados originais.

Quando se faz uma adubação foliar, é importante a época na qual se aplica, pois segundo Ritchie (1998) trabalhos com a cultura do feijoeiro mostraram que a cultura somente respondia a adubação foliar com nitrogênio quando aplicadas aos 30 e 45 dias após a emergência das plantas, verificando-se um maior número de grãos desenvolvidos por vagem, enquanto as aplicações mais tardias aumentavam o peso das sementes mas não proporcionavam maiores produtividades. Portanto, o melhor resultado obtido no início do desenvolvimento da planta é muito bom, pois é no estágio V3 que se pode realizar correções na cobertura da plantação e aumentar a produtividade da cultura da soja ainda na mesma safra.

As Figuras 36, 37 e 38 mostram o gráfico de dispersão para os dados reais (eixo x) e os dados simulados pela melhor RNA (eixo y) separadas por coleta.

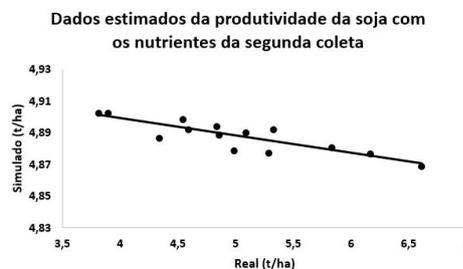


Figura 36 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da primeira coleta de nutrientes da folha.

Fonte: Autoria própria.

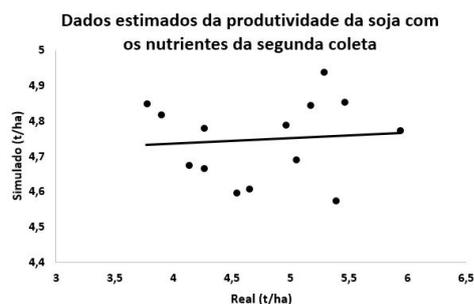


Figura 37 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da segunda coleta de nutrientes da folha.

Fonte: Autoria própria.

Pela análise do coeficiente de determinação, a primeira coleta é a única época que viabiliza a RNA prever a produtividade da soja. Como a primeira coleta deve ser realizada em um estágio bem inicial do desenvolvimento da planta, isto permite ao produtor reverter possíveis produtividades baixas a tempo da colheita.

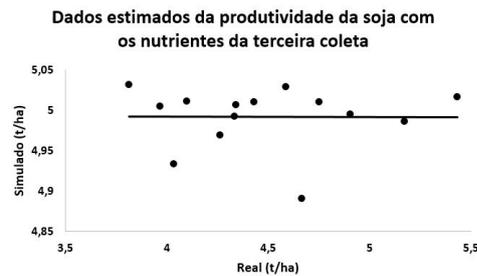


Figura 38 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da terceira coleta de nutrientes da folha.

Fonte: Autoria própria.

O coeficiente de correlação linear de Pearson foi utilizado para medir a relação linear entre os dados reais e os simulados pela RNA. De acordo com a classificação de Callegari e Sidia (2003), a RNA obtida com os dados da primeira coleta tem uma forte correlação linear, apesar de ter sido negativa, o que significa que o modelo de RNA é muito bom, pois a correlação de $-0,86$ é bem próxima de -1 , quando se tem uma correlação linear perfeita negativa. Para a segunda e terceira coleta a correlação linear é classificada como fraca.

A Tabela 11 mostra o valor de cada ponto amostral utilizado no conjunto de teste e o valor simulado em cada RNA para a coleta 1, 2 e 3 dos nutrientes da folha da soja.

Tabela 11 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha.

Coleta 1			Coleta 2			Coleta 3		
Ponto	Real	Simulado	Ponto	Real	Simulado	Ponto	Real	Simulado
B53	4,5390	4,8988	B22	5,1730	4,8450	B61	4,3360	4,9568
B39	5,3280	4,8923	B28	3,8950	4,8176	B69	4,3420	4,9598
B61	4,3360	4,8868	A5	5,0450	4,6919	B31	3,8140	4,9652
B70	4,8580	4,8890	B72	5,3900	4,5748	A39	4,9050	4,9572
B25	4,5890	4,8924	A24	4,9590	4,7891	A17	5,4340	4,9618
B40	4,8340	4,8945	B71	4,2640	4,6672	B71	4,2640	4,9519
A16	6,6090	4,8692	A28	5,2870	4,9384	B24	4,1000	4,9608
B31	3,8140	4,9029	B37	4,2620	4,7798	B52	3,9670	4,9593
B60	4,9870	4,8791	B18	5,4620	4,8545	B36	4,0330	4,9443
A2	5,0850	4,8903	B41	4,1330	4,6751	B68	4,6650	4,9354
A28	5,2870	4,8776	B57	4,5420	4,5977	B59	4,7500	4,9606
B7	6,1640	4,8769	B34	3,7730	4,8486	B22	5,1730	4,9555
A11	5,8290	4,8809	B17	5,9370	4,7739	B67	4,4300	4,9606
B28	3,8950	4,9026	B65	4,6510	4,6078	B32	4,5880	4,9646

Fonte: Autoria própria.

Para fazer um comparativo dos dados reais e os simulados pelas RNAs e verificar o

quanto as amostras são homogêneas entre os dois conjuntos de dados, realizou-se o teste não paramétrico de Kruskal-Wallis (PONTES; CORRENTE, 2001), devido aos dados brutos não apresentarem normalidade (Tabela 12).

Tabela 12 – Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor RNA de cada coleta dos nutrientes da folha da soja.

Kruskal-Wallis		
	H	P-valor
C1	0	1
C2	0,019	0,8903
C3	10,0513	0,0015

Fonte: Autoria própria.

Na Tabela 12, observa-se que as RNAs da coleta 1 e 2 dos nutrientes da soja obteve p-valor $> 0,05$, o que significa que o teste de Kruskal-Wallis aceita a hipótese de homogeneidade entre os dados reais e simulados ao nível de 5% de significância. Já o modelo de RNA da terceira coleta dos nutrientes da folha da soja, rejeita a hipótese de que os dados reais e simulados vem de populações semelhantes.

O valor crítico é 3,565 (valor analisado na tabela do qui-quadrado utilizada para testes de Kruskal-Wallis ao nível de 5% de confiança), o que implica que se o valor de H for menor que o valor crítico, os dados reais e simulados não possuem diferenças significativas, como ocorre na coleta 1 e 2, porém na coleta 3, além de os dados não serem semelhantes, o valor crítico comprova que há diferenças significativas entre os grupos analisados (grupo dos dados reais e outro dos dados simulados), pois o valor de H é maior que o valor crítico.

Como o melhor modelo de RNA foi encontrado para a primeira coleta dos nutrientes da folha da soja, a Figura 39 representa os erros quadráticos médios de treinamento e validação da rede até o término de sua execução, onde o erro de validação passou a aumentar. Pode se perceber que o erro de treinamento começou em 0,1 e abaixou para 0,02, sendo sempre menor que o erro de validação da rede conforme a quantidade de épocas foi aumentando.

A Figura 40 mostra a topologia do modelo da rede neural artificial criada para coleta 1 dos nutrientes da soja e os pesos sinápticos de cada conexão entre os neurônios no momento em que a rede parou seu processo de treinamento.

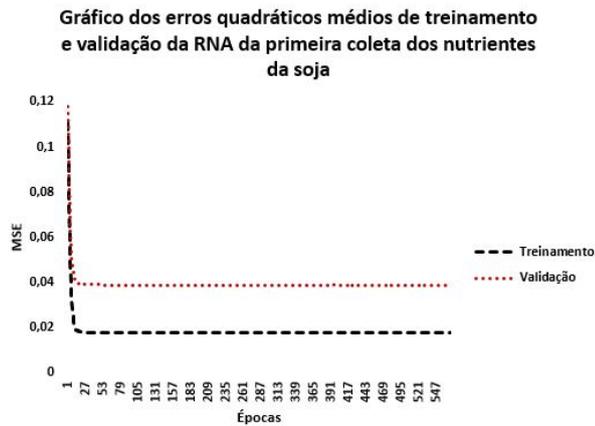


Figura 39 – Gráfico de comparação dos erros quadráticos médios de treinamento e validação até o término da execução da RNA treinada com os nutrientes da primeira coleta da folha da soja.

Fonte: Autoria própria.

Realizou-se testes para as três coletas das áreas A e B, definindo os mesmos pontos amostrais para os conjuntos de treinamento, validação e teste. Porém os resultados da Tabela 10 foram melhores na primeira coleta e piores na segunda e terceira coletas comparando-os aos da Tabela 13, na qual os coeficientes de determinação demonstram que os dados simulados na primeira coleta explicam até 8% da variação dos dados reais e os coeficientes de correlação mostram fraca correlação dos dados reais e simulados em qualquer uma das coletas. E se comparar as coletas 2 e 3 dos modelos de RNAs da Tabela 10, os resultados foram piores, pois a correlação foi menor e os coeficientes de determinação também.

Tabela 13 – Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Topologia	Épocas	MSE	R^2	$R\rho$
1	1x5_1x2_1x1	276	0,0211	0,0831	0,2882
2	1x5_1x2_1x1	166	0,0212	0,0512	0,2262
3	1x5_2x445_1x1	65819	0,0203	0,0376	-0,1938

Fonte: Autoria própria.

A Tabela 14 mostra os pontos amostrais utilizados no conjunto de teste e os valores preditos pelos modelos de RNA da primeira, segunda e terceira coleta dos nutrientes da soja.

Mesmo utilizando RNAs com até três camadas intermediárias de 1 a 500 neurônios em cada uma, não houve nenhum modelo entre os melhores que precisou de três camadas intermediárias, e os modelos utilizaram pouquíssimos neurônios com exceção do modelo obtido na terceira coleta da Tabela 13, que utilizou 445 neurônios em cada camada intermediária.

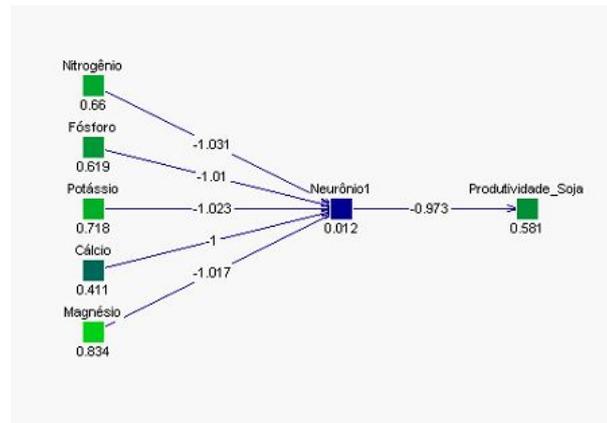


Figura 40 – Ilustração do melhor modelo de RNA com os sinais de entrada, os pesos sinápticos, o limiar de ativação, potencial de ativação e função de ativação, utilizando os dados da primeira coleta da folha da soja.

Fonte: Autoria própria.

Foram realizados modelos de RNAs para as áreas A e B separadas, para verificar se o desempenho dos modelos estava sendo prejudicado pelas diferenças entre as áreas, desde diferenças nutricionais a diferenças de relevo e plantio. A Tabela 15 mostra os resultados obtidos pelas melhores RNAs de cada coleta dos nutrientes da soja somente para a área A.

Analisando os resultados obtidos da área A separadamente da área B, foram melhores do que os resultados das áreas juntas, pois em todas as coletas o coeficiente de determinação apresentou 31 a 61% da variância dos dados reais pelo modelo de RNA, enquanto os testes das duas áreas juntas (utilizando os mesmos pontos amostrais para os conjuntos de treinamento, validação e teste) apresentaram menos de 1% da variação dos dados reais da produtividade da soja. As topologias dos melhores modelos foram com apenas uma camada intermediária e com muito poucos neurônios (apenas um na segunda e terceira coleta). A correlação dos dados preditos e reais pelo modelo de rede neural da segunda e terceira coleta (com melhores coeficientes de determinação) é forte e correlação moderada negativa para a primeira coleta dos nutrientes da folha da soja. O erro varia de 0,01 a 0,02.

Na Tabela 16 verifica-se os valores de produtividade dos pontos amostrais utilizado no conjunto de teste e os valores preditos.

A Figura 41 mostra a análise de regressão para o melhor modelo da área A, ou seja, para a RNA da segunda coleta dos nutrientes da soja. Apesar de os dados reais e simulados terem sido muito semelhantes, o teste de Kruskal-Wallis mostrou que as diferenças entre os dados reais e simulados da segunda coleta são significativas e o p-valor $< 0,05$ demonstrou a heterogeneidade dos dados ao nível de 5% de significância.

Para a área B, os resultados não foram tão bons quantos na área A (Tabela 17). As

Tabela 14 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta (utilizando os mesmos pontos amostrais para treinamento, validação e teste para as três coletas).

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
B58	4,8252	4,8299	4,8321	4,8885
B45	4,3831	4,8577	4,8356	4,8736
A2	5,0850	4,8424	4,7875	4,8731
A22	5,9200	4,8582	4,8517	4,8499
B6	5,7561	4,8361	4,8253	4,8697
A29	5,0071	4,8404	4,8362	4,7752
A8	5,5190	4,8384	4,8035	4,8446
B67	4,4302	4,8211	4,7701	4,8703
B65	4,6510	4,8414	4,7566	4,8625
B54	4,9249	4,8245	4,7979	4,8494
A27	4,3099	4,8352	4,8503	4,8634

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

Tabela 15 – Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja para a área A e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Topologia	Épocas	MSE	R^2	$R\rho$
1	1x5_1x3_1x1	18806	0,0186	0,3195	-0,5652
2	1x5_1x1_1x1	59	0,0200	0,6147	0,7840
3	1x5_1x1_1x1	57	0,0199	0,5220	0,7225

Fonte: Autoria própria.

Tabela 16 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha somente na área A.

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
A14	5,6720	4,9972	4,9583	4,9672
A22	5,9275	5,0513	4,9647	4,9641
A26	5,2902	5,0389	4,9632	4,9677
A10	6,2021	4,9862	4,9697	4,9717
A4	5,0707	5,0439	4,9309	4,9546

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

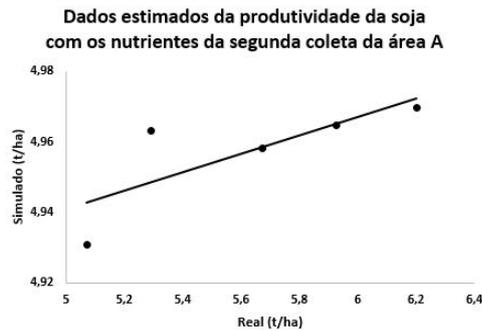


Figura 41 – Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da segunda coleta de nutrientes da folha da soja na área A.

Fonte: Autoria própria.

topologias obtiveram maiores quantidades de neurônios nas camadas intermediárias (coletas 1 e 3). Observa-se pelo coeficiente de determinação que as coletas 1 e 3 tiveram modelos que explicaram de 3 a 4% dos dados reais e mesmo a coleta 2 que obteve um coeficiente de determinação maior, o modelo explicou 36% dos dados. Os erros quadráticos médios foram um pouco maiores também que os erros da área A e as correlações variaram de fraca a moderada correlação dos dados reais e preditos.

Tabela 17 – Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja para a área B e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Topologia	Épocas	MSE	R^2	$R\rho$
1	1x5_2x369_1x1	550597	0,0437	0,0415	-0,2037
2	1x5_2x1_1x1	4464	0,0407	0,3625	0,6021
3	1x5_1x214_1x1	171255	0,0457	0,0375	-0,1937

Fonte: Autoria própria.

A Tabela 18 mostra os pontos amostrais utilizados no conjunto de testes para a área B e os respectivos valores reais e preditos pelos melhores modelos de RNAs.

Portanto, os melhores resultados foram obtidos dos dados das duas áreas juntas na primeira coleta dos nutrientes da soja e dos dados das áreas separadas na segunda coleta dos nutrientes da folha da soja na área A.

O melhor resultado obtido foi com os nutrientes da primeira coleta dos dados, ou seja, dos dados do estágio V3, enquanto os demais estádios tiveram desempenhos bastante inferiores. Os estádios V6 e R1 podem ter sido influenciados pela menor quantidade de nutrientes restante na folha, pois segundo Mundstock e Thomas (2005) e Ritchie (1998), no estágio V6 e nos estádios reprodutivos, se inicia um período de rápido e constante acúmulo

Tabela 18 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs de cada coleta dos nutrientes da folha somente na área B.

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
B40	4,8385	4,7579	4,8531	4,7209
B57	4,5469	4,7226	4,6715	4,4494
B73	4,8614	4,8521	4,9931	4,7621
B68	4,6697	4,8845	5,0589	4,1159
B38	4,6258	4,8565	4,4668	4,6572
B21	4,0475	4,8637	4,5561	4,8221

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

diário das taxas de matéria seca e de nutrientes pela planta, que se inicia nas partes vegetativas (folhas, hastes, pecíolos e raízes), deslocando-se gradualmente para as vagens e sementes em formação, enquanto as partes vegetativas finalizam o seu desenvolvimento.

Segundo Rosolem (2002) quanto mais velha menor a absorção de nutrientes da folha e os nutrientes são transmitidos para as flores e vagens da planta.

Com relação a adubação na cobertura da plantação, Faquin (2002) comenta da importância em realizar a análise foliar para complementar os resultados da análise do solo e que o uso da análise da planta pode representar grande economia de fertilizantes e ganhos na produção.

A análise foliar pode ser usada para recomendação de adubação, visando alcançar a máxima produtividade e a melhor qualidade de grãos, porém segundo Borkert et al. (1994) basta que as folhas da soja sejam colhidas no início da floração e pleno florescimento para utilizar a análise foliar como base para adubação. Mas pelos resultados obtidos, a RNA obteve melhores previsões no início do estágio vegetativo da planta V3, sendo possível realizar a adubação e corrigir deficiências e excessos de nutrientes para aumentar a produtividade na safra atual.

4.3 DESEMPENHO DAS MÁQUINAS DE SUPORTE DE VETORES

Utilizou-se os parâmetros padrão de treinamento do algoritmo SMOReg da API do software WEKA, como a constante C de complexidade = 1,0, N = 2 para não normalizar os dados, pois já estão normalizados, o parâmetro epsilon - L = 0,001, o número de semente

aleatório - $W = 1$, o epsilon de erro de arredondamento - $P = 1,0E-12$, o parâmetro de tolerância para verificar o critério de paragem - $T = 0,001$.

Utilizou-se 60% dos dados para treinamento, 30% para validação e 10% para teste. Executou-se o algoritmo de treinamento com o kernel polinomial e obteve-se os resultados da Tabela 19 para cada coleta dos nutrientes da soja.

Tabela 19 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja (pontos amostrais aleatórios para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	MSE	R^2	$R\rho$
1	Polinomial	0,1552	0,2342	-0,4839
2	Polinomial	0,0164	0,0227	0,1506
3	Polinomial	0,0253	0,1969	-0,4437

Fonte: Autoria própria.

O menor erro MSE é da SVM da segunda coleta dos nutrientes da soja, enquanto o melhor coeficiente de determinação é da SVM da primeira coleta, que também tem o melhor índice de correlação dos dados reais e simulados. Portanto, a SVM da primeira coleta foi o melhor modelo de predição dos dados dentre as três SVMs das três coletas, sendo capaz de explicar 23% da variação dos dados. A terceira coleta obteve um modelo capaz de explicar aproximadamente 20% da variação dos dados originais, enquanto que a segunda coleta obteve a pior SVM, capaz de explicar 1% apenas.

Pela classificação de Callegari e Sidia (2003), a correlação dos dados reais e simulados da SVM implementada com os dados da primeira e terceira coleta dos nutrientes da folha da soja é uma correlação linear moderada. Enquanto a SVM da segunda coleta possui correlação fraca.

Para melhor visualização dos resultados, as Figuras 42, 43, 44 mostram a inclinação da reta da regressão linear dos dados reais e simulados da produtividade da soja em toneladas por hectare.

Na Tabela 20, verifica-se que as SVMs das três coletas dos nutrientes da soja obtiveram p-valor $> 0,05$, o que significa que o teste de Kruskal-Wallis aceita a hipótese de homogeneidade entre os dados reais e simulados ao nível de 5% de significância.

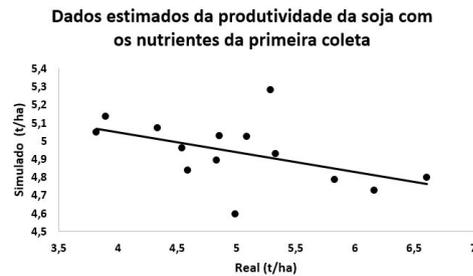


Figura 42 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da primeira coleta de nutrientes da folha.

Fonte: Autoria própria.

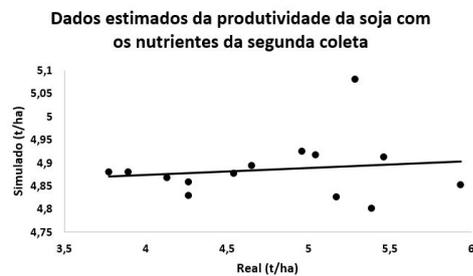


Figura 43 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da segunda coleta de nutrientes da folha.

Fonte: Autoria própria.

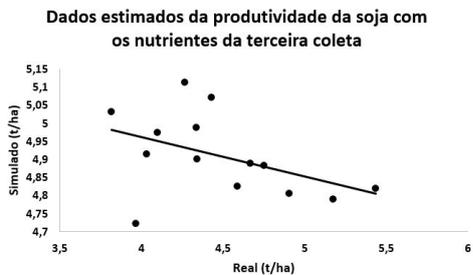


Figura 44 – Resultado dos valores estimados da produtividade da soja com a SVM obtida com os dados da terceira coleta de nutrientes da folha.

Fonte: Autoria própria.

O modelo de SVM da terceira coleta dos nutrientes da folha da soja por pouco não rejeita a hipótese de que os dados reais e simulados vem de populações semelhantes. O valor crítico foi analisado na tabela do qui-quadrado utilizada para testes de Kruskal-Wallis, e corresponde a 3,565, deste modo como valor de H é menor que o valor crítico nas coletas 1 e 2, não há diferenças significativas entre os grupos (dados reais e simulados), e para coleta 3, apesar de os dados poderem ser considerados de uma mesma população, as diferenças entre

eles são significativas (Tabela 20).

Tabela 20 – Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja.

Kruskal-Wallis		
	H	P-valor
C1	0,0021	0,9633
C2	0,0084	0,9267
C3	7,8557	0,0051

Fonte: Autoria própria.

Na tabela 21, verifica-se os valores reais de cada ponto amostral utilizado no conjunto de teste e o valor simulado pela SVM de cada coleta dos nutrientes da folha da soja.

Tabela 21 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta dos nutrientes da folha.

Coleta 1			Coleta 2			Coleta 3		
Ponto	Real	Simulado	Ponto	Real	Simulado	Ponto	Real	Simulado
B53	4,5390	4,9591	B22	5,1730	4,8254	B61	4,3360	4,9873
B39	5,3280	4,9294	B28	3,8950	4,8790	B69	4,3420	4,9006
B61	4,3360	5,0720	A5	5,0450	4,9162	B31	3,8140	5,0307
B70	4,8580	5,0258	B72	5,3900	4,8010	A39	4,9050	4,8044
B25	4,5890	4,8382	A24	4,9590	4,9248	A17	5,4340	4,8194
B40	4,8340	4,8910	B71	4,2640	4,8586	B71	4,2640	5,1131
A16	6,6090	4,7957	A28	5,2870	5,0802	B24	4,1000	4,9732
B31	3,8140	5,0484	B37	4,2620	4,8289	B52	3,9670	4,7220
B60	4,9870	4,5944	B18	5,4620	4,9111	B36	4,0330	4,9139
A2	5,0850	5,0219	B41	4,1330	4,8671	B68	4,6650	4,8873
A28	5,2870	5,2790	B57	4,5420	4,8758	B59	4,7500	4,8813
B7	6,1640	4,7271	B34	3,7730	4,8802	B22	5,1730	4,7895
A11	5,8290	4,7867	B17	5,9370	4,8519	B67	4,4300	5,0704
B28	3,8950	5,1338	B65	4,6510	4,8927	B32	4,5880	4,8250

Fonte: Autoria própria.

Realizou-se testes para as três coletas das áreas A e B, definindo os mesmos pontos amostrais para os conjuntos de treinamento, validação e teste. Porém os resultados da Tabela 22 foram piores comparado a aqueles da Tabela 19, pois os coeficientes de determinação demonstram que os dados simulados explicam até 1% da variação dos dados reais e os coeficientes de correlação mostram fraca correlação dos dados reais e simulados em qualquer uma das coletas.

Também realizou-se testes para as coletas 1, 2 e 3 separando os pontos amostrais das áreas A e B, ou seja, obteve-se resultados de SVMs para as coletas 1, 2 e 3 para a área A e B

Tabela 22 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	MSE	R^2	$R\rho$
1	Polinomial	0,0262	0,0072	0,0849
2	Polinomial	0,0272	0,0178	0,1334
3	Polinomial	0,0275	0,0723	-0,2689

Fonte: Autoria própria.

separadas.

Os resultados foram satisfatórios para a área A (Tabela 23), pois obteve-se modelos de SVMs que explicaram de 23% a 58% da variação total dos dados reais da produtividade da soja, com destaque para a segunda coleta, a qual obteve maior coeficiente de correlação, mostrando uma correlação negativa mas classificada como forte de acordo com classificação de Callegari e Sidia (2003).

Tabela 23 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja somente na área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	MSE	R^2	$R\rho$
1	Polinomial	0,0262	0,3759	-0,6131
2	Polinomial	0,0293	0,5868	-0,7660
3	Polinomial	0,0247	0,2359	0,4856

Fonte: Autoria própria.

A Tabela 24 mostra os valores reais de cada ponto utilizado para o conjunto de teste e o valor estimado de cada SVMs das coletas 1, 2 e 3.

Tabela 24 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta somente da área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
A14	5,6720	4,9344	5,1069	4,8334
A22	5,9275	5,2252	4,9154	4,8144
A26	5,2902	4,9389	4,9319	5,0075
A10	6,2021	4,8124	4,8161	5,0794
A4	5,0707	5,5852	5,2654	4,7409

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

Os modelos tiveram bons desempenhos, mas pelo teste de Kruskal-Wallis rejeita-se a

hipótese de os dados reais e simulados serem homogêneos ao nível de 5% de significância e analisando o valor de H que é maior que o valor crítico do qui-quadrado (3,565), as diferenças entre os grupos (dados reais e simulados) são significativas. Portanto, os dados reais e simulados não podem ser considerados provenientes de uma mesma população (Tabela 25).

Tabela 25 – Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja da área A.

Kruskal-Wallis		
	H	P-valor
C1	3,9381	0,0472
C2	4,8109	0,0282
C3	5,7709	0,0162

Fonte: Autoria própria.

Para a área B, os resultados não foram tão bons quanto da área A (Tabela 26). Obteve-se coeficientes de determinação mais baixos, variando de 0,08 a 0,3799 e coeficientes de correlação de fraca a forte correlação. Assim como na área A, a área B obteve melhores resultados com o modelo de SVM para a segunda coleta, onde obteve-se uma semelhança dos dados reais e simulados de aproximadamente 38% e uma forte correlação entre eles.

Tabela 26 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs de cada coleta dos nutrientes da folha da soja somente na área B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	MSE	R^2	$R\rho$
1	Polinomial	0,0211	0,229	0,4785
2	Polinomial	0,0487	0,3799	0,6163
3	Polinomial	0,0545	0,0824	0,2871

Fonte: Autoria própria.

Na Tabela 27 estão os pontos amostrais utilizados no conjunto de teste e seus respectivos valores simulados por cada modelo das coletas 1, 2 e 3 da área B.

Analisando os resultados dos testes de Kruskal-Wallis para a área B (Tabela 28), todas as coletas apresentaram homogeneidade nos dados reais e simulados ao nível de 5% de significância e verifica-se por meio do valor de H, que as diferenças entre os conjuntos de dados reais e simulados não são significativas.

Tabela 27 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores SVMs de cada coleta somente da área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
B40	4,8339	4,7896	5,0451	4,7533
B57	4,5419	4,7741	4,7256	4,8605
B73	4,8569	5,4113	5,1689	4,6628
B68	4,6650	4,7055	5,3562	4,8180
B38	4,6211	4,8857	4,2422	4,8249
B21	4,0421	4,7408	4,4011	4,6421

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

Tabela 28 – Testes de homocedasticidade e semelhança de população entre os dados reais e simulados pela melhor SVM de cada coleta dos nutrientes da folha da soja da área B.

Kruskal-Wallis		
	H	P-valor
C1	2,5641	0,1093
C2	0,923	0,3366
C3	0,923	0,3366

Fonte: Autoria própria.

Portanto, apesar de a área A ter tido melhores resultados nos modelos, não foi possível constatar que os dados reais e simulados eram provenientes de uma mesma população. Na área B, os modelos não foram tão bons, mas apresentaram homogeneidade. Deste modo, a segunda coleta dos nutrientes da soja na área B, teve um desempenho geral melhor para o modelo de predição (SVM) da produtividade criado.

4.3.1 Comparação dos resultados obtidos pelas RNAs e SVMs

Os modelos de RNAs obtidos dos dados divididos em conjuntos de treinamento, validação e teste com pontos escolhidos aleatoriamente obteve melhores desempenhos tanto nas RNAs quanto nas SVMs na primeira coleta dos nutrientes da folha da soja.

Já para as demais coletas não se obteve bons resultados. Analisando os conjuntos de treinamento, validação e teste divididos com os mesmos pontos amostrais nas três coletas, os resultados foram ruins em ambas as coletas e técnicas de inteligência artificial.

Quando se verifica os resultados dos modelos de RNA e SVM somente dos pontos amostrais da área A, os melhores modelos de RNA e SVM tiveram coeficientes de determinação parecidos nas coletas 1 e 2. Na coleta 3, o coeficiente de determinação é maior no modelo de RNA, o qual explica 52% dos dados reais, enquanto o melhor modelo de SVM da terceira coleta explicou 23% dos dados reais. A RNA também obteve melhor desempenho na segunda coleta em relação a SVM, analisando as técnicas de avaliação dos métodos de predição. E na primeira coleta, acontece o inverso.

Nos modelos obtidos somente dos dados da área B, a SVM obteve melhores desempenhos em relação a RNA de acordo com os parâmetros de avaliação utilizados neste trabalho.

Portanto, a RNA teve melhores desempenhos em alguns casos, enquanto a SVM obteve melhores desempenhos em outros. No entanto, o melhor modelo de todos os modelos treinados, desde os modelos das áreas juntas aos modelos das áreas separadas, a RNA com somente um neurônio na única camada intermediária com os dados da primeira coleta dos nutrientes da soja coletados das duas áreas amostrais deste estudo, obteve o maior coeficiente de determinação, explicando 74% dos dados reais, possuindo a maior correlação dos dados reais e preditos.

O segundo melhor modelo obtido, foi com os dados da área A da segunda coleta dos nutrientes da soja, obtido também com uma RNA, coincidentemente com um neurônio e uma camada intermediária explicando 61% dos dados reais e a segunda maior correlação dos dados preditos e reais da produtividade da soja dentre todos os modelos de RNAs e SVMs treinados.

A área A obteve, em geral, resultados melhores em comparação a área B, pois a área A teve seu plantio realizado em um único dia, enquanto a área B, por ser uma área maior é dividida em duas partes do talhão para ser plantada em dias diferentes. Desta forma, os dados das folhas e dos grãos foram coletados no mesmo dia, quando na verdade uma parte da área B deveria ser coletada dias após a parte que foi plantada antes, devido a não completa maturação da planta. Essas coletas da área B foram realizadas no mesmo dia devido a dificuldade de reunir muitas pessoas em vários dias para ajudar, pois cada um possui compromissos e conciliar datas em um grande grupo de pessoas se torna difícil. Deste modo, os dados ficam com uma certa discrepância na área B e isto influenciou nos resultados. Portanto, a área A possui resultados mais relevantes que a área B.

Em desempenho geral a RNA obteve melhores modelos de predição da produtividade da soja com os testes realizados.

4.4 SELEÇÃO DE ATRIBUTOS

Utilizou-se a técnica CfsSubsetEval do software WEKA, em que CFS remete à seleção de atributos baseada na correlação. Este algoritmo acopla uma fórmula de avaliação com uma medida de correlação e uma estratégia de busca heurística (HALL, 1999).

Segundo Hall (1999) a Correlation based Feature Selection (CFS) consegue identificar facilmente atributos redundantes e quais os atributos mais relevantes de uma base de dados.

Os atributos obtidos com esta técnica de seleção de atributos podem ser visualizados nas Tabelas 29 e 30.

Tabela 29 – Atributos selecionados pela técnica CfsSubsetEval para as coletas 1, 2 e 3 das áreas A e B juntas.

Coleta	Atributos selecionados
1	Ca, Mg
2	P, Ca
3	K, Mg

Fonte: Autoria própria.

Tabela 30 – Atributos selecionados pela técnica CfsSubsetEval para as coletas 1, 2 e 3 das áreas A e B separadas.

Área	Coleta	Atributos selecionados
A	1	K, Ca, Mg
	2	N
	3	P, Ca, Mg
B	1	N, Ca, Mg
	2	N, P, Ca
	3	K, Ca, Mg

Fonte: Autoria própria.

Buscou-se outras formas de selecionar os atributos mais adequados para treinamento das RNAs e SVMs como a matriz de correlação de Pearson. O coeficiente de correlação de Pearson é utilizado para medir o quanto uma variável se correlaciona a outra. Porém, na matriz de correlação é necessário analisar o p-valor de cada correlação da matriz, pois o coeficiente de correlação só é válido se o valor de p for muito menor que 1 e muito maior que -1, ou seja, com um nível de significância de 5% (SENSUSEA et al., 2015).

Portanto, realizou-se a matriz de correlação de Pearson para obter o quanto as variáveis de Nitrogênio, Potássio, Fósforo, Cálcio e Magnésio se correlacionavam com os dados da produtividade, para que se utilizasse nos modelos de RNAs e SVMs, apenas as melhores variáveis para estimar a produtividade da soja.

A matriz de correlação de Pearson foi realizada para as coletas 1, 2 e 3 com as duas áreas amostrais juntas e também separadas. No entanto, nenhuma das correlações teve o p-valor menor que 0,05 de significância. Sendo assim, não foi possível realizar testes com as variáveis correlacionadas com a produtividade da soja utilizando o coeficiente de correlação de Pearson.

4.4.1 Resultado dos melhores modelos de RNAs obtidos

Treinou-se 9000 modelos de redes neurais artificiais, porém, os atributos selecionados pela técnica CfsSubsetEval não tiveram bons desempenhos nos modelos de RNAs treinados. Conforme a Tabela 31 verifica-se que o coeficiente de determinação variou de 0,8 a 11% e a correlação dos dados simulados e reais foi fraca para ambas as coletas dos nutrientes selecionados pela técnica.

Tabela 31 – Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja nas áreas A e B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta)

Coleta	Topologia	Atributos	Épocas	MSE	R^2	$R\rho$
1	1x2_2x3_1x1	Ca_Mg	16665	0,0181	0,0293	-0,171
2	1x2_1x1_1x1	P_Ca	343	0,0206	0,1149	-0,339
3	1x2_1x94_1x1	K_Mg	6648	0,0203	0,0089	0,0943

Fonte: Autoria própria.

A Tabela 32 mostra a tendência das RNAs em obter saídas no caso da primeira coleta entre 4,8 a 5,1 ton/ha, e na coleta 2, as saídas preditas variam entre 4,7 a 4,8 ton/ha, enquanto a terceira coleta varia de 4,6 a 4,8 ton/ha.

Para a área A, os atributos selecionados na primeira coleta (K, Ca e Mg) não obtiveram bons desempenhos se comparar ao resultado do melhor modelo com os cinco atributos da folha coletados (Tabela 33). A segunda coleta obteve um resultado bom com apenas um nutriente na entrada da RNA (N), porém não foi tão bom quanto o modelo de RNA com os cinco nutrientes

Tabela 32 – Valor real da produtividade da soja para os pontos amostrais utilizados no conjunto de testes e o valor simulado pelas melhores RNAs utilizando somente os atributos indicados pela técnica CfsSubsetEval de cada coleta das áreas A e B.

Ponto	Real	Simulado_C1	Simulado_C2	Simulado_C3
B58	4,8252	4,9712	4,8427	4,6024
B45	4,3831	4,8694	4,8445	4,7447
A2	5,0850	4,8732	4,8645	4,8037
A22	5,9200	5,0192	4,8544	4,8328
B6	5,7561	4,7868	4,7793	4,7907
A29	5,0071	4,9935	4,8516	4,8881
A8	5,5190	4,8918	4,8672	4,8530
B67	4,4302	4,8030	4,8591	4,8206
B65	4,6510	5,0562	4,8550	4,8167
B54	4,9249	4,9196	4,8612	4,8227
A27	4,3099	5,0779	4,8535	4,8885

C1-Coleta 1; C2-Coleta 2; C3-Coleta 3.

Fonte: Autoria própria.

(N, P, K, Ca e Mg).

Tabela 33 – Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta)

Coleta	Topologia	Atributos	Épocas	MSE	R^2	$R\rho$
1	1x3_1x2_1x1	K_Ca_Mg	10185	0,0185	0,2271	-0,4766
2	1x1_1x13_1x1	N	6589	0,0178	0,5081	-0,7128
3	1x3_1x6_1x1	P_Ca_Mg	4155	0,0186	0,0014	0,0374

Fonte: Autoria própria.

Na Tabela 34 verifica-se os resultados obtidos dos melhores modelos de RNAs para as coletas dos nutrientes da folha da soja na área B com os atributos selecionados pela técnica CfsSubsetEval. Assim como na área A, a melhor RNA obtida foi na segunda coleta dos nutrientes da folha da soja na área B. O coeficiente de determinação da segunda coleta mostra que os dados reais foram explicados pelos dados preditos em até 56%, apresentando uma forte correlação entre eles. Analisando o número de épocas, a primeira e terceira coleta realizaram muito mais ajustes em relação a segunda, além de terem utilizado maior quantidade de neurônios na camada intermediária, e não apresentaram resultados melhores e nem erros quadráticos médios menores.

Portanto, para as áreas A e B tanto separadas quanto unidas, os melhores resultados foram na segunda coleta dos nutrientes da folha da soja.

Tabela 34 – Topologias, quantidade de épocas até o momento de parada da execução da rede, e seus respectivos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores RNAs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta)

Coleta	Topologia	Atributos	Épocas	MSE	R^2	$R\rho$
1	1x3_1x208_1x1	N_Ca_Mg	122468	0,0457	0,0603	-0,2456
2	1x3_1x1_1x1	N_P_Ca	1592	0,0424	0,5669	0,7529
3	1x3_1x238_1x1	K_Ca_Mg	58453	0,0447	0,0479	-0,2189

Fonte: Autoria própria.

4.4.2 Resultado dos melhores modelos de SVMs obtidos

Os resultados obtidos com as áreas juntas tiveram desempenhos piores com os atributos selecionados. No entanto, os resultados das áreas A e B separadas foram diferentes, sendo melhores os resultados da área B com os atributos selecionados. Para a área A os resultados das SVMs utilizando somente os atributos indicados pela técnica CfsSubsetEval foram um pouco inferiores aos resultados das SVMs que utilizaram todos os atributos coletados da folha da soja, porém a diferença foi mínima.

A Tabela 35 mostra os resultados dos modelos de SVMs obtidos para as coletas 1, 2 e 3 utilizando os atributos Ca e Mg, P e Ca, K e Mg, respectivamente.

Os coeficientes de determinação não foram satisfatórios, pois os dados simulados explicaram no máximo 10% dos dados reais possuindo correlação fraca em ambas as coletas dos respectivos nutrientes da soja selecionados pela técnica de seleção de atributos.

Tabela 35 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja nas áreas A e B(mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	Atributos	MSE	R^2	$R\rho$
1	Polinomial	Ca_Mg	0,027	0,102	-0,3193
2	Polinomial	P_Ca	0,0275	0,0147	0,1214
3	Polinomial	K_Mg	0,027	0,0265	-0,1627

Fonte: Autoria própria.

Na Tabela 36 verifica-se os resultados referentes a área A, onde a segunda coleta dos nutrientes, obteve melhor desempenho utilizando apenas o Nitrogênio para a predição da produtividade. O modelo da segunda coleta do Nitrogênio obteve 52% de semelhança dos dados

simulados e reais, apresentando uma forte correlação negativa entre os dados reais e simulados.

Tabela 36 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área A (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	Atributos	MSE	R^2	$R\rho$
1	Polinomial	K_Ca_Mg	0,026	0,3488	-0,5905
2	Polinomial	N	0,0261	0,5201	-0,7211
3	Polinomial	P_Ca_Mg	0,0246	0,1435	0,3788

Fonte: Autoria própria.

A área B, obteve os resultados apresentados na Tabela 37, onde os coeficientes de determinação mostram que os dados simulados apresentaram a variação dos dados reais de 33% a 58%. Assim como na área A, os melhores resultados na área B se apresentaram na segunda coleta dos nutrientes da soja utilizando apenas Nitrogênio, Potássio e Cálcio. As correlações dos dados reais e simulados variaram de moderada a forte correlação.

Tabela 37 – Resultado dos erros MSE, coeficiente de determinação e de correlação de Pearson para as melhores SVMs treinadas somente com os atributos indicados pela técnica de seleção de atributos CfsSubsetEval de cada coleta dos nutrientes da folha da soja na área B (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta).

Coleta	Kernel	Atributos	MSE	R^2	$R\rho$
1	Polinomial	N_Ca_Mg	0,0501	0,3358	0,5794
2	Polinomial	N_P_Ca	0,0488	0,5866	0,7658
3	Polinomial	K_Ca_Mg	0,0551	0,39	0,6245

Fonte: Autoria própria.

4.4.3 Comparação dos resultados obtidos utilizando a técnica CfsSubsetEval

Nos modelos de RNAs e SVMs utilizando os respectivos atributos selecionados pela técnica de seleção de atributos em cada uma das coletas (1, 2 e 3), os resultados não foram os esperados, pois houve baixa correlação dos dados reais e os preditos e os valores dos dados preditos ficaram muito distantes dos valores reais da produtividade da soja.

Para a área A, os resultados foram melhores, obtendo-se desempenhos bastante semelhantes nos modelos de RNA e SVM. Na área B, os modelos de RNAs não tiveram bons

resultados, com exceção da segunda coleta, que na SVM também apresentou mais de 50% da variação dos dados reais.

5 CONSIDERAÇÕES FINAIS

Neste capítulo são resumidas as principais considerações finais acerca do trabalho, bem como suas contribuições econômicas, sociais e ambientais. Também é abordado possíveis trabalhos futuros para melhorar e dar continuidade ao objetivo de utilizar técnicas de inteligência artificial para prever a produtividade de uma cultura.

5.1 CONCLUSÃO

Este trabalho mostrou que é possível obter um modelo de predição da produtividade da soja com os nutrientes da folha, pois uma rede neural artificial do tipo Perceptron Multicamadas utilizando o algoritmo *Backpropagation* conseguiu explicar 74% da variação dos dados reais, utilizando para isto, 5 neurônios na camada de entrada, um neurônio em uma única camada intermediária e um neurônio na camada de saída. Com este modelo é possível realizar a correção de possíveis fatores de baixa na produtividade na cobertura da plantação no início de seu desenvolvimento, a tempo de melhorar a produtividade para a safra atual da adubação.

A correlação dos dados reais e preditos pelo melhor modelo da técnica de inteligência artificial obtido ficou classificada como forte ($R\rho = 0,86$), e o critério de parada antecipada baseado na validação teve o ponto considerado como mínimo local na 67ª época.

A aplicação da técnica de seleção de atributos diminuiu a quantidade de atributos como entrada para as RNAs e SVMs, mas na maioria dos resultados não obteve melhora com relação aos modelos que utilizaram todos os atributos. Porém, os modelos de SVM utilizando os dados da área B, considerando a diminuição da quantidade de variáveis de entrada, teve resultados muito semelhantes ao modelo que utilizou os cinco atributos coletados da folha da soja. Se analisar o lado econômico da pesquisa, se torna mais barato prever a produtividade da soja, pois seriam necessários menos atributos para aplicar na técnica de inteligência artificial.

Na SVM que utilizou somente os atributos selecionados, a segunda coleta dos

nutrientes da folha obteve um modelo que explicou 58% da variação dos dados reais utilizando os nutrientes de Nitrogênio, Fósforo e Cálcio, obtendo o mesmo desempenho verificado na segunda coleta dos dados da área A pela SVM utilizando além dos atributos de Nitrogênio, Fósforo e Cálcio, os nutrientes de Potássio e Magnésio, ou seja, neste caso, a seleção de atributos deveria ser utilizada. No entanto, o melhor modelo de todos os modelos treinados, não se obteve com a redução de atributos de entrada nos conjuntos de treinamento.

Desta forma, conclui-se que pode ser utilizado o modelo de RNA que prediziu 74% dos dados coletados neste trabalho para prever a produtividade da soja, ou como alternativa de custo mais baixo, pode ser utilizado o modelo de SVM com os atributos selecionados pela técnica CfsSubsetEval para que a aplicação da técnica de inteligência artificial seja mais acessível e possa de uma maneira também eficiente verificar futuras possíveis baixas produtividades para realizar correções necessárias ao aumento da produção e rendimentos do produtor somente nos locais críticos.

5.2 TRABALHOS FUTUROS

Verificou-se ao longo desta pesquisa, que há poucos trabalhos de inteligência artificial voltados para estimar a produtividade de uma cultura. Como apresentado na conclusão deste trabalho, as técnicas de inteligência artificial são muito promissoras para auxiliar na predição da produtividade, e é necessário continuar a explorá-las buscando alterar seus parâmetros, realizar outras formas de avaliação e utilizar outras heurísticas de otimização do treinamento das técnicas.

No entanto, não se deve pesquisar somente as técnicas utilizadas neste trabalho, pois há muitas técnicas além de redes neurais artificiais e máquinas de suporte de vetores que podem ser estudadas para buscar resultados ainda melhores.

Grandes melhorias nestas pesquisas podem ser realizadas utilizando outras características da planta e do solo, como os componentes químicos, físicos e de relevo do solo, o tamanho das folhas, a quantidade de grãos por planta em diferentes estágios, entre outras.

A seleção de atributos pode auxiliar no desempenho das técnicas preditivas e há muitas técnicas de seleção de atributos que podem ser aplicadas em trabalhos futuros para predição de produtividades não só da cultura da soja, mas também de outras. As técnicas que levam em consideração a dependência espacial dos atributos podem melhorar os resultados preditos, pois

há muitas características que estão associadas a localização de certas variáveis e é importante levá-las em consideração. Também pode ser utilizado a seleção de atributos conforme a correlação dos mesmos, comprovada cientificamente por engenheiros da área agrícola.

Outro interessante trabalho é buscar aplicar as técnicas de Inteligência Artificial utilizando os dados separados por zonas de manejo, pois as zonas de manejo são sub-regiões de um talhão divididas de acordo com as suas características similares do solo. Desta forma, os resultados podem ser mais satisfatórios, pois uma técnica de predição receberá dados de entrada e de produtividade de uma região com a mesma capacidade produtiva.

Portanto, há inúmeras pesquisas que são muito interessantes e importantes na agricultura de precisão que devem ser incentivadas a serem concretizadas, pois obter técnicas capazes de prever a produtividade de uma cultura resulta em inúmeras vantagens para a economia, sociedade e meio ambiente.

REFERÊNCIAS

- ACOSTA, J. A. de A.; LEMAINSKI, C. L.; GABRIEL, O.; SANTI, R.; LONDERO, G.; BUSATO, M. Constante evolução. **Revista Cultivar**, n. 131, p. 12–16, 2010.
- ANTUNIASSI, U. R.; BAIO, F. H. R.; SHARP, T. C. Agricultura de precisão. In: Congresso Brasileiro do Algodão, 6., 2007, Uberlândia. **Anais...** Uberlândia, 2007. p. 11–21.
- BERNARDI, A. C. de C.; NAIME, J. de M.; RESENDE Álvaro Vilela de; BASSOI, L. H.; INAMASU, R. Y. **Agricultura de Precisão: Resultados de um novo olhar**. Brasília: Embrapa, 2014. 596 p.
- BOCCA, F. F. **Produtividade de cana-de-açúcar: caracterização dos contextos de decisão e utilização de técnicas de mineração de dados para modelagem**. 98 f. Dissertação (Mestrado em Engenharia Agrícola) — Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas (UNESP), Campinas, 2014.
- BORKERT, C. M.; YORINORI, J. T.; CORRÊA-FERREIRA, B. S.; ALMEIDA Álvaro M.; FERREIRA, L. P.; SFREDO, G. J. **Seja o doutor da sua Soja**. 1. ed. Piracicaba, 1994. 17 p.
- CALLEGARI, J.; SIDIA, M. **Bioestatística: princípios e aplicações**. Porto Alegre: Artemed, 2003. 255 p.
- CARDON, A.; MULLER, D. N. **Introdução às Redes Neurais Artificiais**. 31 f. Dissertação (Mestrado em Ciência da Computação) — Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1994.
- CITRIX SYSTEMS INCORPORATION. **ABOUT XENSERVR**. 2015. Disponível em: <<http://xenserver.org/about-xenserver-open-source.html>>. Acesso em: 08 de maio de 2016.
- COSTA, E. M. da; SILVA, H. F.; RIBEIRO, P. R. de A. Matéria orgânica do solo e o seu papel na manutenção e produtividade dos sistemas agrícolas. **Revista Enciclopédia Biosfera**, v. 9, n. 17, p. 1842–1860, 2013.
- CYBENKO, G. **Approximations by superpositions of sigmoidal functions**. 2. ed. [S.l.]: Mathematics of Control, Signals and Systems, 1989. 303-314 p.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. 2. ed. New York: Wiley, 2001. 680 p.
- EPSTEIN, E.; BLOOM, A. **Nutrição Mineral de Plantas: Princípios e perspectivas**. 2. ed. Londrina: Planta, 2006. 404 p.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. de. **Inteligência Artificial: Uma abordagem de aprendizado de máquina**. 1. ed. Rio de Janeiro: LTC, 2011. 378 p.
- FAQUIN, V. **Diagnose do estado nutricional das plantas**. 1. ed. Lavras: UFLA/FAEPE, 2002. 77 p.

- FERNANDES, A. M. da R. **Inteligência Artificial: noções gerais**. 1. ed. Florianópolis: Visual Books, 2003. 160 p.
- FILHO, L. S. M.; MOL, A. A.; ROCHA, R. Desenvolvimento de ferramenta computacional para auxílio ao projeto de gemas lapidadas. **Revista Escola de Minas**, v. 58, n. 4, p. 367–373, 2005.
- GIANLUPPI, V.; GIANLUPPI, D.; JÚNIOR, A. L. M.; ZILLI, J. Édson; NECHET, K. de L.; BARBOSA, G. ferreira; MATTIONI, J. A. M. **Correção e manutenção da fertilidade do solo. Cultivo de Soja no Cerrado de Roraima**. Embrapa Roraima, 2009. Disponível em: <<https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Soja/CultivodeSojanoCerradodeRoraima/adubacao.htm>>. Acesso em: 08 de abril de 2016.
- HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA data mining software: An update. **SIGKDD Explorations**, v. 11, n. 1, p. 10–18, 2009.
- HALL, M. A. **Correlation-based Feature Selection for Machine Learning**. Tese (Doutorado em Filosofia) — University of Waikato, New Zealand, 1999.
- HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 900 p.
- LEAL, A. J. F.; MIGUEL, E. P.; BAIO, F. H. R.; NEVES, D. de C.; LEAL, U. A. S. Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo. **Bragantia**, Campinas, v. 74, n. 4, p. 436–444, 2015.
- LORENA, A. C.; CARVAHO, A. C. P. L. F. de. **Introdução às Máquinas de Vetores Suporte**. São Carlos, Abril 2003. 58 p.
- LORENA, A. C.; CARVALHO, A. C. P. L. F. de. Uma introdução às support vector machines. **RITA**, v. 14, n. 2, p. 43–67, 2007.
- MACHADO, A. B. M. **Neuroanatomia Funcional**. 2. ed. São Paulo: Atheneu, 2003. 360 p.
- MAMIER, G.; WIELAND, J. **Stuttgart Neural Network Simulator Batchman User Manual**. 1995. Disponível em: <<http://ssdi.di.fct.unl.pt/~nmm/scripts/snns/batchman.pdf>>. Acesso em: 08 de maio de 2016.
- MENEZES, P. L. de. **Desenvolvimento de modelo em redes neurais para simulação da precipitação na irrigação por aspersão**. Tese (Doutorado em Engenharia Agrícola) — Universidade Federal de Campina Grande, Campina Grande, 2014.
- MERCER, J. Functions of positive and negative type, and their connection with the theory of integral equations. **Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**, The Royal Society, v. 209, n. 441-458, p. 415–446, 1909.
- MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO. **Agricultura de Precisão**. 2. ed. Brasília, 2011. 36 p.
- MITCHELL, T. M. **Machine Learning**. 1. ed. New York: McGraw-Hill, Inc., 1997. 432 p.

MOLIN, J. P. **Agricultura de Precisão: o gerenciamento da variabilidade**. 1. ed. Piracicaba: o autor, 2001. 81 p.

MOLIN, J. P. **Agricultura de Precisão: situação atual e perspectivas**. Piracicaba, 2003. 89-98 p.

MONDO, V. H. V.; JUNIOR, F. G. G.; PINTO, T. L. F.; MARCHI, J. L. de; MOTOMIYA, A. V. de A.; MOLIN, J. P.; CICERO, S. M. Spatial variability of soil fertility and its relationship with seed physiological potential in a soybean production area. **Revista Brasileira de Sementes**, v. 34, p. 193–201, 2012.

MOREIRA, M. G. **Soja – Análise da Conjuntura Agropecuária**. Curitiba, 2015. 16 p.

MUNDSTOCK, C. M.; THOMAS, A. L. **Soja: Fatores que afetam o crescimento e o rendimento de grãos**. 1. ed. Porto Alegre: Evangraf, 2005. 31 p.

NOVAIS, R. F. de; V., V. H. A.; BARROS, N. F. D.; FONTES, R. L.; CANTARUTTI, R. B.; NEVES, J. C. L. **Fertilidade do solo**. 1. ed. Viçosa: Sociedade Brasileira de Ciência do Solo, 2007. 1017 p.

Organização para a Cooperação e Desenvolvimento Econômico. **Perspectivas Ambientais da OCDE para 2030**. 2008. Disponível em: <<http://www.oecd.org/env/indicators-modelling-outlooks/40220494.pdf>>. Acesso em: 10 de março de 2016.

PIMENTEL, F. G.; GARCIA, H. C. **Estatística aplicada a experimentos agrônômicos e florestais: exposição com exemplos e orientações para o uso de aplicativos**. Piracicaba: Fealq, 2002. 309 p.

PONTES, A. C. F.; CORRENTE, J. E. Comparações múltiplas não-paramétricas para o delineamento com um fator de classificação simples. **Revista de Matemática e Estatística**, v. 19, p. 179–197, 2001.

RITCHIE, S. W. **Como a planta de soja se desenvolve**. Piracicaba, 1998. 23 p.

ROSOLEM, C. A. **Recomendação e aplicação de nutrientes via foliar**. 1. ed. Lavras: UFLA/FAEPE, 2002. 98 p.

RUB, G. Data mining of agricultural yield data: A comparison of regression models. In: **Proceedings of the 9th Industrial Conference on Advances in Data Mining, Applications and Theoretical Aspects**. Heidelberg: Springer-Verlag, 2009. (ICDM '09), p. 24–37.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A modern approach**. 3. ed. New Jersey: Prentice Hall, 2009. 932 p.

SCHÖLKOPF, B.; SMOLA, A. J. **Learning with Kernels: Support vector machines, regularization, optimization, and beyond**. Cambridge: MIT Press, 2001. 632 p.

SEDIYAMA, T. **Tecnologias de produção e usos da soja**. Londrina: Mecenasa, 2009. 314 p.

SENSUSEA, D. I.; CAHYANINGSIHA, E.; WIBOWOA, W. C. Identifying knowledge management process of indonesian government human capital management using analytical hierarchy process and pearson correlation analysis. **Procedia Computer Science**, v. 72, p. 233–243, 2015.

SILVA, I. N. D.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático**. 1. ed. São Paulo: Artliber, 2010. 399 p.

SILVA, L. N. de C. **Engenharia Imunológica: Desenvolvimento e Aplicação de Ferramentas Computacionais Inspiradas em Sistemas Imunológicos Artificiais**. Tese (Doutorado em Engenharia Elétrica) — Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação, Campinas, 2001.

SMOLA, A. J.; SCHOLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, Kluwer Academic Publishers, v. 14, n. 3, p. 199–222, 2004.

SOARES, F. C.; RUSSI, A. D. R. J. L.; PEITER, M. X.; PARIZI, A. R. C. Uso de rede neural artificial para predição da produção na cultura do feijoeiro. In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA - CONBEA, 42., 2014, Campo Grande. **Anais...** Campo Grande: Associação Brasileira de Engenharia Agrícola, 2014. p. 7.

SOCIEDADE BRASILEIRA DE CIÊNCIA DO SOLO. **Manual de Adubação e Calagem para os estados do Rio Grande do Sul e Santa Catarina**. 10. ed. Porto Alegre, 2004. 394 p.

SOLAB - Equipamentos para laboratório. **SL-31 - Moinho de facas tipo Willey**. 2016. Disponível em: <<http://www.solabcientifica.com.br/exibir2.php?id2=217&id=Moinhos&orca=74e0f483d97855d08910a3ac17f2f815>>. Acesso em: 17 de maio de 2016.

SU, C. T.; TONG, L. I.; LEOU, C. M. Combination of time series and neural network for reliability forecasting modeling. **Journal of the Chinese Institute of Industrial Engineer**, Beijing, v. 14, n. 4, p. 419–429, 1997.

TEDESCO, M. J.; GIANELLO, C.; BISSANI, C. A.; BOHNEN, H.; VOLKWEISS, S. J. **Análise de solo, plantas e outros materiais**. 2. ed. Porto Alegre, 1995. 174 p.

TUBINGEN, U. of. **JavaNNS: Java Neural Network Simulator**. 2016. Disponível em: <<http://www.ra.cs.uni-tuebingen.de/software/JavaNNS/>>. Acesso em: 08 de maio de 2016.

VITTI, G. C.; TREVISAN, W. **Manejo de macro e micronutrientes para alta produtividade da soja**. Piracicaba, 2000. 16 p.

WHELAN, B.; TAYLOR, J. Precision agriculture for grain production systems. **International Journal of Digital Earth**, Csiro, v. 6, n. 6, p. 610–612, 2013.

WOLLENHAUPT, N. C.; MULLA, D. J.; CROWFORD, G. C. Soil sampling and interpolation techniques for mapping spatial variability of soil properties. **The Site-Specific Management for Agricultural Systems**, p. 19–53, 1997.

XAVIER, A. E. Uma função ativação para redes neurais artificiais mais flexível e poderosa e mais rápida. **Learning and Nonlinear Models - Revista da Sociedade Brasileira de Redes Neurais (SBRN)**, v. 3, n. 1, p. 1–7, 2005.

YADAV, N.; YADAV, A.; KUMAR, M. **An Introduction to Neural Network Methods for Differential Equations**. 1. ed. Springer Netherland: Springer, 2015. 114 p.

ZHANG, F.-S.; YAMASAKI, S. ya; NANZYU, M. Waste ashes for use in agricultural production: I. liming effect, contents of plant nutrients and chemical characteristics of some metals. **The Science of the Total Environment**, v. 284, n. 1-3, p. 215–255, 2002.