

**UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E BIOLÓGICAS
EMBRAPA MANDIOCA E FRUTICULTURA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS GENÉTICOS VEGETAIS
CURSO DE MESTRADO**

**INTELIGÊNCIA ARTIFICIAL APLICADA EM ESTUDOS DE MODELAGEM E
DIVERGÊNCIA GENÉTICA EM GENÓTIPOS DE INHAME (*Dioscorea
rotundata*)**

INGRID GIOVANNA VIEIRA SANTOS

CRUZ DAS ALMAS – BA
FEVEREIRO/2022

**INTELIGÊNCIA ARTIFICIAL APLICADA EM ESTUDOS DE MODELAGEM E
DIVERGÊNCIA GENÉTICA EM GENÓTIPOS DE INHAME (*Dioscorea
rotundata*)**

INGRID GIOVANNA VIEIRA SANTOS

Bacharelado em Ciências Biológicas

Universidade Federal do Vale do São Francisco (UNIVASF), 2019

Dissertação apresentada ao Colegiado do Programa de Pós-Graduação em Recursos Genéticos Vegetais da Universidade Federal do Recôncavo da Bahia e Embrapa Mandioca e Fruticultura, como requisito parcial para obtenção do grau de Mestre em Recursos Genéticos Vegetais.

Orientador: Prof. Dr. Ricardo Franco Cunha Moreira

Coorientador: Prof. Dr. Liniker Fernandes da Silva

CRUZ DAS ALMAS – BA

2022

FICHA CATALOGRÁFICA

S237i	<p>Santos, Ingrid Giovanna Vieira. Inteligência artificial aplicada em estudos de modelagem e divergência genética em genótipos de inhame (<i>Dioscorea rotundata</i>) / Ingrid Giovanna Vieira Santos._ Cruz das Almas, Bahia, 2022. 82f.; il.</p> <p>Dissertação (Mestrado) – Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, Mestrado em Recursos Genéticos Vegetais.</p> <p>Orientador: Prof. Dr. Ricardo Franco Cunha Moreira. Coorientador: Prof. Dr. Liniker Fernandes da Silva.</p> <p>1.Inhame – Inteligência artificial. 2.Inhame – Melhoramento genético – Análise. I.Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas. II.Título.</p> <p style="text-align: right;">CDD: 635.16</p>
-------	---

Ficha elaborada pela Biblioteca Central de Cruz das Almas - UFRB.
Responsável pela Elaboração - Antonio Marcos Sarmiento das Chagas (Bibliotecário - CRB5 / 1615).
(os dados para catalogação foram enviados pela usuária via formulário eletrônico).

UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA
CENTRO DE CIÊNCIAS AGRÁRIAS, AMBIENTAIS E BIOLÓGICAS
EMBRAPA MANDIOCA E FRUTICULTURA
PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS GENÉTICOS VEGETAIS
CURSO DE MESTRADO

INTELIGÊNCIA ARTIFICIAL APLICADA EM ESTUDOS DE MODELAGEM E
DIVERGÊNCIA GENÉTICA EM GENÓTIPOS DE INHAME (*Dioscorea*
***rotundata*)**

Comissão Examinadora da Defesa de Dissertação

Ingrid Giovanna Vieira Santos

Aprovada em: ____ de _____ de ____

Prof. Dr. Ricardo Franco Cunha Moreira
Universidade Federal do Recôncavo da Bahia (UFRB)
Orientador

Prof. Dr. Sebastião de Oliveira e Silva
Universidade Federal do Recôncavo da Bahia (UFRB)
Examinador Interno

Prof. Dra. Bianca Machado Campos Trindade
Universidade Federal Rural do Rio de Janeiro (UFRRJ)
Examinadora Externa

“Confie no Senhor de todo o seu coração e não se apoie em seu próprio entendimento; reconheça o Senhor em todos os seus caminhos, e ele endireitará as suas veredas.” (Provérbios 3:5,6)

AGRADECIMENTOS

Agradeço primeiramente a Deus, sem Ele jamais chegaria até aqui. Obrigada Senhor por conduzir os meus passos e me livrar de todo mal. A Sua proteção me fortalece e me traz paz. Sei que há muito que percorrer, mas com a Sua presença eu sei que consigo enfrentar qualquer desafio. Não tenho palavras para agradecer tamanha gratidão. Muito obrigada!

Aos meus pais, José Roberto e Vera Vieira, por todo o amor que têm por mim. Sempre acreditaram no meu potencial e me incentivaram nas horas mais difíceis desta caminhada, principalmente nos momentos em que o estresse e o cansaço tomaram conta de mim. Essa vitória é nossa. Vocês são meu porto seguro. Amo vocês!

Aos meus irmãos, Ícaro Roberto, Ítalo Ramon, Iago Ronan e lally Rayane, por todo o cuidado, carinho e amizade. Obrigada pelos três pimpolhos da minha vida: Maria Luíza, João Lucas e Kaleb.

Ao meu amor, Erick Oliveira, que está sempre ao meu lado, me apoiando e me mostrando o quão eu sou capaz. Obrigada por todo o cuidado e proteção que tem por mim. Você é muito especial em minha vida. Sem dúvida, você tornou essa caminhada mais leve.

A toda a minha família que sempre me apoiou. Agradeço por tudo!

À família merenda que me recebeu de braços abertos. Obrigada por tanto carinho que tem por mim.

Ao meu orientador, Prof. Ricardo Franco, e ao meu coorientador, Prof. Liniker Fernandes, pelos ensinamentos, confiança e paciência. Obrigada por todo apoio prestado durante a realização desse trabalho. É gratificante tê-los em minha formação acadêmica. As universidades estão precisando de mais seres humanos como vocês. Gratidão!

À minha amiga, Geisa Matos, pela amizade, conselhos, incentivos e companheirismo desde o início do mestrado. Obrigada pelo carinho, cuidado e por compartilhar momentos inesquecíveis durante essa caminhada. Desejo-te o

melhor e que sua trajetória seja sempre vitoriosa, pois é uma menina dedicada e determinada. Que Deus te abençoe abundantemente! Tenho muito orgulho de você. Sempre estará guardada em meu coração. Saudades!

Aos meus amigos, Ian e Lucas, pela ajuda, incentivo e palavras de apoio durante essa trajetória.

Aos meus professores que foram muito importantes para minha formação. Obrigada por cada ensinamento.

À Universidade Federal do Recôncavo da Bahia (UFRB) por disponibilizar toda infraestrutura necessária para a minha formação acadêmica.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Enfim, obrigada a todos que contribuíram na realização desse trabalho e na minha trajetória acadêmica.

**INTELIGÊNCIA ARTIFICIAL APLICADA EM ESTUDOS DE MODELAGEM E
DIVERGÊNCIA GENÉTICA EM GENÓTIPOS DE INHAME (*Dioscorea
rotundata*)**

Autora: Ingrid Giovanna Vieira Santos

Orientador: Prof. Dr. Ricardo Franco Cunha Moreira

Coorientador: Prof. Dr. Liniker Fernandes da Silva

RESUMO: O objetivo deste estudo foi avaliar a predição do peso do tubérculo do inhame mediante a análise de Redes Neurais Artificiais, bem como estimar e organizar a divergência genética dos genótipos de inhame com base na análise dos Mapas Auto-Organizáveis de Kohonen. Os genótipos foram coletados nos municípios de Cruz das Almas, São Felipe, São Félix e Maragogipe, sendo posteriormente caracterizados e avaliados mediante os descritores preconizados pelo Instituto Internacional de Recursos Genéticos Vegetais e Instituto Internacional de Agricultura Tropical (IPGRI/IITA). Os resultados indicaram que as Redes Neurais Artificiais se apresentaram como uma ferramenta satisfatória na predição do peso do tubérculo do inhame e a análise do Mapa Auto-Organizável de Kohonen foi eficaz para distinguir a dissimilaridade dos genótipos de inhame na Região do Recôncavo Baiano, com vista a subsidiar programas de melhoramento genético, uma vez que podem facilitar decisões antecipadas por parte do agricultor e/ou melhorista, além de possibilitar ao produtor fazer uma estimativa das projeções para a safra e prever aspectos de financiamento antes da colheita.

PALAVRAS-CHAVE: Redes Neurais Artificiais; Mapa auto-organizável de Kohonen; Agricultura; Melhoramento genético vegetal.

ARTIFICIAL INTELLIGENCE APPLIED IN MODELING STUDIES AND GENETIC DIVERGENCE IN YAM GENOTYPES (*Dioscorea rotundata*)

Author: Ingrid Giovanna Vieira Santos

Advisor: Prof. Dr. Ricardo Franco Cunha Moreira

Co-Advisor: Prof. Dr. Liniker Fernandes da Silva

ABSTRACT: The aim of this study was to evaluate the prediction of yam tuber weight through the analysis of Artificial Neural Networks, as well as to estimate and organize the genetic divergence of yam genotypes based on the analysis of Kohonen's Self-Organizing Maps. The genotypes were collected in the municipalities of Cruz das Almas, São Felipe, São Félix and Maragogipe, being later characterized and evaluated using the descriptors recommended by the International Institute of Plant Genetic Resources and the International Institute of Tropical Agriculture (IPGRI/IITA). The results indicated that the Artificial Neural Networks were presented as a satisfactory tool in the prediction of the yam tuber weight and the analysis of the Kohonen Self-Organizing Map was effective to distinguish the dissimilarity of the yam genotypes in the Recôncavo Baiano Region, with a view to subsidize genetic improvement programs, since they can facilitate early decisions on the part of the farmer and/or breeder, in addition to enabling the producer to estimate the projections for the harvest and foresee financing aspects before the harvest.

KEYWORDS: Artificial Neural Networks; Kohonen's Self-Organizing Maps; Agriculture; Plant genetic improvement.

LISTA DE FIGURAS

REVISÃO DE LITERATURA

Figura 1. Distribuição geográfica das espécies <i>D. rotundata</i> , <i>D. abyssinica</i> e <i>D. praeheensis</i>	18
Figura 2. Representação de um neurônio biológico	22
Figura 3. Modelo de neurônio artificial, segundo McCulloch-Pitts. Onde x_1, x_2, \dots, x_m são as entradas da rede; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos, ou pesos sinápticos, associados a cada entrada; b_k é o termo bias; u_k é a combinação linear dos sinais de entrada; $\varphi(.)$ é a função de ativação e y_k é a saída do neurônio	22
Figura 4. Rede Neural Artificial de múltiplas camadas	24
Figura 5. Representação gráfica das funções de ativação mais utilizadas: (a) função limiar; (b) função linear; (c) função sigmoideal e (d) função tangente hiperbólica	26
Figura 6. Mecanismo de aprendizado supervisionado	28
Figura 7. Mecanismo de aprendizado não supervisionado	28
Figura 8. Representação de um Mapa Auto-Organizável de Kohonen	29
Figura 9. Representação gráfica do número de ciclos para o treinamento da rede	31

CAPÍTULO I

Figura 1. Representação do neurônio artificial dentro de uma rede do tipo Perceptron de Múltiplas Camadas ou <i>Multilayer Perceptron</i> (MLP)	47
Figura 2. Gráfico de resíduos em função do peso estimado do tubérculo do inhamo das cinco melhores redes neurais artificiais treinadas na etapa de validação	52/53

CAPÍTULO II

Figura 1. Arquitetura de uma rede neural SOM bidimensional com topologia do tipo hexagonal	71
Figura 2. Organização de similaridade dos genótipos de inhamo (<i>Dioscorea rotundata</i>) a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5), onde baseia-se na intensidade da coloração da área preenchida do círculo para indicar a concentração de genótipos em cada neurônio	73

Figura 3. Matriz de distância unificada entre neurônios vizinhos dos genótipos de inhame (*Dioscorea rotundata*) a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5) 76

LISTA DE TABELAS

CAPÍTULO I

Tabela 1. Relação dos genótipos de inhame (*Dioscorea rotundata*) caracterizados na Região do Recôncavo Baiano com seus respectivos códigos, série, procedência (povoado e município) e número 45

Tabela 2. Estatística das variáveis mensuradas nos genótipos de inhame (*Dioscorea rotundata*) caracterizados na Região do Recôncavo Baiano 50

Tabela 3. Redes Neurais Artificiais treinadas para estimativa do peso do tubérculo do inhame (*Dioscorea rotundata*) em função de caracteres agrônômicos na Região do Recôncavo Baiano 51

CAPÍTULO II

Tabela 1. Relação dos genótipos de inhame (*Dioscorea rotundata*) caracterizados na Região do Recôncavo Baiano com seus respectivos códigos, série, procedência (povoado e município) e número 69

Tabela 2. Agrupamento em neurônios dos vinte povoados da Região do Recôncavo Baiano, nos quais foram cultivados os genótipos de inhame (*Dioscorea rotundata*), a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5) 74

SUMÁRIO

1	INTRODUÇÃO	15
2	REVISÃO DE LITERATURA	16
2.1	Aspectos botânicos e taxonômicos do gênero <i>Dioscorea</i>	16
2.2	Origem e dispersão do gênero <i>Dioscorea</i>	17
2.3	Importância socioeconômica e nutricional do gênero <i>Dioscorea</i>	19
2.4	Redes Neurais Artificiais	21
2.4.1	<i>Funções de ativação</i>	24
2.4.2	<i>Paradigmas de aprendizagem</i>	27
2.4.3	<i>Vantagens e dificuldades da implementação das Redes Neurais Artificiais</i>	30
2.4.4	<i>Aplicação das Redes Neurais Artificiais no melhoramento</i> ...	31
	REFERÊNCIAS BIBLIOGRÁFICAS	34

CAPÍTULO I

	REDES NEURAS ARTIFICIAIS PARA A PREDIÇÃO DO PESO DO TUBÉRCULO DE INHAME (<i>Dioscorea rotundata</i>)	40
1	INTRODUÇÃO	43
2	MATERIAL E MÉTODOS	44
2.1	Área de estudo	44
2.2	Redes Neurais Artificiais	46
3	RESULTADOS	49
4	DISCUSSÃO	54
4.1	Importância da viabilização da cultura de inhame na busca da produtividade	54
4.2	Desenvolvimento e utilização da modelagem na agricultura	56
4.3	Aprimoramento da modelagem para predição na cultura do inhame	58
5	CONCLUSÃO	58
	REFERÊNCIAS BIBLIOGRÁFICAS	60

CAPÍTULO II

REDES NEURAIS ARTIFICIAIS NA ANÁLISE DE DIVERGÊNCIA GENÉTICA EM INHAME (<i>Dioscorea rotundata</i>)	64
1 INTRODUÇÃO	67
2 MATERIAL E MÉTODOS	68
2.1 Área de estudo	68
2.2 Redes Neurais Artificiais	70
3 RESULTADOS	72
4 DISCUSSÃO	76
5 CONCLUSÃO	79
REFERÊNCIAS BIBLIOGRÁFICAS	80

1. INTRODUÇÃO

O inhame é um tubérculo monocotiledôneo pertencente à família Dioscoreaceae e ao gênero *Dioscorea*, apresentando mais de 600 espécies cultivadas nas regiões tropicais, subtropicais e temperadas de todo mundo (EPPING; LAIBACH, 2020). Dentre as espécies que apresentam uma maior relevância econômica destacam-se a *Dioscorea cayennensis*, *D. alata*, *D. trifida*, *D. esculenta* e *D. rotundata* (PADHAN; PANDA, 2020).

O inhame é uma espécie cultivada em todas as regiões brasileiras; entretanto a Região Nordeste é a maior produtora, concentrando-se 90% de toda a produção de inhame do Brasil, principalmente os estados da Paraíba, Pernambuco, Alagoas, Bahia, Maranhão e Piauí, por constituir um bom negócio agrícola em função do alto consumo pela população (IBGE, 2021). A importância do seu cultivo ocorre devido ao fato de ser uma das principais fontes nutricionais de milhões de pessoas (RAMAN et al., 2014). Além da alimentação humana, destaca-se também pela utilização como matéria-prima em inúmeros produtos industriais (MESQUITA, 2002).

A aplicação da inteligência artificial nas atividades agrícolas visando o melhoramento tornou-se uma ferramenta fundamental para auxiliar na tomada de decisão, bem como a previsão e/ou exploração do potencial produtivo da cultura (SOUSA et al., 2020), uma vez que os produtores de inhame requerem variedades melhoradas que combinem resistência a doenças e pragas e alta produtividade com atributos apreciados pelo consumidor (MIGNOUNA et al., 2007; OBIDIEGWU et al., 2009).

Então, a partir de um banco de dados de experimentos bem conduzidos que caracterizam acessos em condições edafoclimáticas, genótipos podem ser selecionados atendendo aos interesses do produtor e/ou melhorista de forma mais eficiente mediante a análise de Redes Neurais Artificiais (RNA) (SILVA et al., 2020).

Segundo Braga, Carvalho e Ludermir (2007), as RNA são métodos heurísticos que vem sendo cada vez mais utilizadas e têm permitido, com o avanço da tecnologia e do entendimento da neurociência, a criação de modelos

de neurônios artificiais que se assemelham aos neurônios biológicos, os quais se conectam formando sinapses artificiais. Essas conexões as tornam capazes de análises de diversas situações, aprendizagem, reconhecimento de padrões e generalização.

Assim, o presente trabalho teve como objetivo avaliar a predição do peso do tubérculo do inhame bem como estimar e organizar a divergência genética dos genótipos da cultura mediante a análise de Redes Neurais Artificiais.

2. REVISÃO DE LITERATURA

2.1 Aspectos botânicos e taxonômicos do gênero *Dioscorea*

O inhame pertence ao gênero *Dioscorea*, o maior e mais importante gênero da família Dioscoreaceae, apresentando mais de 600 espécies (AMANZE et al., 2011; SIQUEIRA et al., 2014). Dentre as principais espécies cultivadas para a utilização no consumo humano, podem-se citar: *Dioscorea cayennensis*, *D. alata*, *D. trifida*, *D. esculenta* e *D. rotundata* (PADHAN; PANDA, 2020).

O inhame é um tubérculo monocotiledôneo, anual e/ou perene, em geral são plantas trepadeiras quando encontram apoio, mas também há espécies eretas e herbáceas (ATIENO et al., 2020). Apresentam o sistema reprodutivo por alogamia e propagação vegetativa por intermédio exclusivo de tubérculos-sementes, que são pequenos tubérculos ou pedaços com aproximadamente 200 g (MONDO et al., 2020).

O caule é aéreo volúvel, desde delgados a robustos, podendo ou não apresentar espinhos (ANOKYE; TETTEH; OTOO, 2014). Os órgãos de reserva são considerados como rizóforos subterrâneos, com ou sem porções espessadas. Além disso, algumas espécies do gênero *Dioscorea* apresentam a peculiaridade de formar órgãos de reserva aéreos nas axilas foliares, denominados de bulbilhos aéreos, os quais acumulam água e nutrientes após a floração (SIQUEIRA, 2009).

As folhas possuem grandes variações morfológicas, geralmente são alternas, opostas ou espiraladas, compostas ou trifoliadas, lobadas ou não, e

pecioladas em forma de coração ou seta. A inflorescência desenvolve-se na axila das folhas superiores, em geral como panícula ou espiga (IPGRI/IITA, 1997; PEDRALLI, 1999; ATIENO et al., 2020). As flores são pequenas, actinomorfas, trímeras e geralmente unissexuais (IPGRI/IITA, 1997), podendo apresentar coloração branca, amarela ou púrpura (ATIENO et al., 2020). As flores masculinas possuem odor adocicado e grãos de pólen viscoso fortemente aderido à antera, além de conter perigônio com seis peças em dois verticilos, apresentando seis estames com anteras férteis, distribuídas em um só verticilo. Enquanto as flores femininas são maiores, apresentam ovário ínfero, tricarpelar, trilocular, geralmente com muitos óvulos e alguns nectários septais (IPGRI/IITA, 1997). Na sua maioria, o inhame é uma planta dioica, embora existam plantas monoicas (AKORODA, 1983; ZOUNDJIHEKPON et al., 1997).

Os frutos são do tipo cápsulas trialadas, bagas ou drupas. As sementes podem ser aladas ou não, reticuladas ou lisas, com tamanhos variados possuindo embrião pequeno bem diferenciado e cotilédone lateral imerso no endosperma, o qual contém lipídios e aleurona (IPGRI/IITA, 1997).

2.2 Origem e dispersão do gênero *Dioscorea*

As espécies do gênero *Dioscorea* originaram-se no Sudeste da Ásia (*D. esculenta*, *D. alata* e *D. bulbifera*), Oeste da África (*D. altíssima*, *D. rotundata*, *D. dumetorum* e *D. cayenensis*) e América Tropical (*D. trifida*) (ALEXANDER; COURSEY, 1969), os quais são considerados centros de diversidade e domesticação da cultura (ASIEDU et al., 1997). Apresentando-se uma grande dispersão entre os continentes, sendo encontrada em regiões tropicais, subtropicais e temperadas de todo o mundo (MONTALDO, 1991; COUTO et al., 2018). De acordo com Coursey (1976), houve uma ampla dispersão mundial do gênero *Dioscorea* no final do Período Cretáceo, posteriormente ao longo do Mioceno, apresentando evolução nas diferentes direções do Novo e Velho Mundo, o que provocou a origem de espécies distintas.

A dispersão do gênero *Dioscorea* deu-se no século XVI, durante o Período Colonial, por intermédio dos portugueses e espanhóis, que eram os maiores navegadores da época (COURSEY, 1967). Entretanto, portugueses e espanhóis

relataram que encontraram os povos indígenas cultivando esse vegetal quando chegaram em terras da América (ABRAMO, 1990). Apesar de cultivado desde a antiguidade por povos primitivos, a dispersão de todas as espécies só chegou à civilização com a intensificação do tráfico negreiro (SOUSA; RAIZADA, 2020).

D. rotundata, também conhecida como inhame da Guiné branco, é a espécie mais importante na África Ocidental e Central, devido ao seu valor nutricional e alimentar, além da importância para a cultura do povo africano (OBIDIEGWU; AKPABIO, 2017). Apesar da relevância do inhame, sua origem apresenta algumas controvérsias. Segundo Coursey (1976), esta espécie pode ser um híbrido entre a espécie selvagem adaptada à savana *D. abyssinica* Hochest ex. Knuth e a espécie selvagem adaptada à floresta tropical *D. praehensilis* Benth mediante a análise morfológica. Esse pressuposto é fundamentado por vários autores, devido a sobreposição parcial das distribuições geográficas das duas espécies (Figura 1).

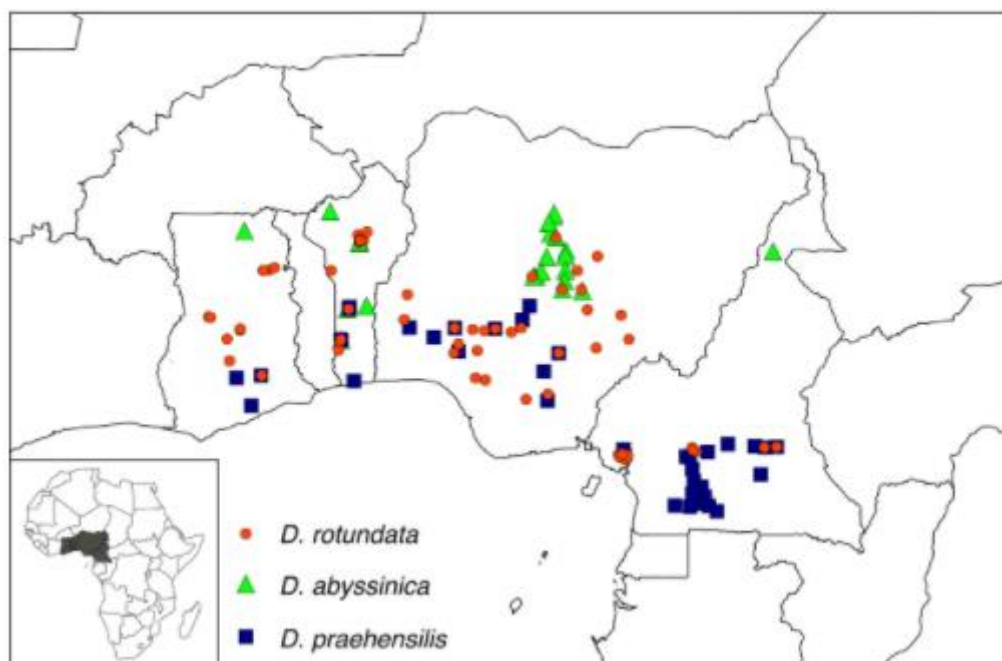


Figura 1. Distribuição geográfica das espécies *D. rotundata*, *D. abyssinica* e *D. praehensilis* na África.

Fonte: SCARCELLI et al. (2019).

Outros autores consideram como sendo uma subespécie da *D. cayenensis*. Entretanto, Ramser et al. (1997) concluíram que as espécies *D. rotundata* e *D. cayenensis* são espécies distintas a partir de análises

taxonômicas e filogenéticas do *Complexo D. rotundata/D. cayenensis* mediante a aplicação de quatro marcadores moleculares, sendo eles: *Random Amplified Polymorphic DNA* (RAPD), Microssatélites, *Random Amplified Microsatellite Polymorphisms* (RAMPO), e por sequências do DNA de cloroplastos. As espécies podem ser fenotipicamente diferenciadas pela coloração da polpa, visto que a da *D. rotundata* é branca e da *D. cayenensis* é amarela, bem como a produção de tubérculos, uma vez que a *D. rotundata* é precoce enquanto a *D. cayenensis* é tardia.

2.3 Importância socioeconômica e nutricional do gênero *Dioscorea*

O inhame é considerado uma cultura de fácil cultivo e manejo por exigir pouco uso de insumos agrícolas e manejo técnico especializado quando comparado com outras culturas (ANDRES; ADEOLUWA; BHULLAR, 2017), o que contribui para sua utilização na agricultura familiar em todas as regiões do Brasil. Portanto, o inhame é uma espécie de fundamental importância para o país, pelo motivo de ser uma das culturas de maior interesse para a agricultura de subsistência e segurança alimentar (SIQUEIRA, 2011).

No entanto, a produção do inhame tem sofrido uma considerável evolução, passando de uma atividade predominantemente de pequenas propriedades rurais para uma atividade importante sob o ponto de vista socioeconômico por gerar empregos diretos e indiretos, uma vez que sua cadeia produtiva envolve outros setores, como: armazenamento, transporte e comercialização (USDA, 2016).

A África Ocidental, especificamente Nigéria, República do Gana, Costa do Marfim, República do Benim e República Togolesa, concentram-se 93,49% do total da produção mundial, desenvolvendo 74,32 milhões de toneladas de túberas no ano de 2019, com uma área plantada de 8.310.702 ha. O Brasil encontra-se entre os principais países produtores de inhame e ocupa a 11^a colocação no ranking mundial em termos de produção. Houve um crescimento na produção brasileira de inhame, que passou de 231,5 mil toneladas em 2000 para 249,5 mil toneladas em 2019, ocorrendo um aumento de cerca de 7% (FAO, 2021).

O inhame é uma espécie cultivada em todas as regiões brasileiras; tendo a Região Nordeste como maior produtora, concentrando-se 90% de toda a produção de inhame do Brasil, principalmente nos estados da Paraíba, Pernambuco, Alagoas, Bahia, Maranhão e Piauí, por constituir um bom negócio agrícola em função do alto consumo pela população. Em seguida, a região que se destaca na produção de inhame é o Sudeste, um importante ingrediente utilizado na culinária mais sofisticada dessa região brasileira (IBGE, 2021).

Devido às condições edafoclimáticas propícias para seu cultivo, além de suas propriedades nutricionais e o consumo demandado pelo mercado interno, a produção de inhame no Nordeste representa uma atividade agrícola promissora (SANTOS, 2002; MENDES; SILVA; FAVERO, 2013). Um outro aspecto favorável é a exportação deste tubérculo, que representa uma importante opção de mercado em função dos preços oferecidos (MENDES; SILVA; FAVERO, 2013).

Por ser um alimento de excelente qualidade nutritiva e energética, rico em vitaminas do complexo B (com altos teores de tiamina, riboflavina, niacina e adermina), teores de vitaminas A, D e C, carboidratos (amido, principalmente), minerais, por exemplo, fósforo, cálcio, ferro, potássio, sódio, magnésio, zinco, manganês, cobre e enxofre, e com propriedades medicinais, além de apresentar baixos teores de gordura e índice glicêmico, conferindo uma melhor proteção contra obesidade e diabetes (BATISTA et al., 2019; KULASINGHEA; RANAWEERA, 2019), o inhame se enquadra em um alimento funcional que deve ter seu consumo estimulado, transpondo assim as barreiras do regionalismo (NORONHA, 2014).

Tanto no mercado interno quanto no externo, as aplicações do inhame se multiplicam, com ampla utilização na forma fermentada (BATISTA et al., 2019), minimamente processada e *in natura* (TRINDADE et al., 2011), podendo ser empregada desde produtos alimentícios até como insumo na fabricação de cosméticos e nas indústrias farmacêuticas, especialmente na síntese de cortisona e de hormônios esteroides (MESQUITA, 2002). O inhame direcionado às indústrias também é utilizado para produção de farinhas, que desempenha uma função fundamental na dieta da população brasileira, uma vez que pode substituir outras farinhas em qualquer receita, como: pães, bolos, panquecas,

pizzas, biscoitos, doces, tortas, entre outras (BRASIL, 2015). É comumente vendido *in natura* em feiras livres e supermercados (ODU; ELENWO; MADUKA, 2019), assim como pode ser encontrado de forma minimamente processado, congelado cru ou pré-cozido, refrigerado e frito (*chips*) (TRINDADE et al., 2011; OMOHIMI et al., 2019).

2.4 Redes Neurais Artificiais

Na agricultura, os produtores buscam atingir uma estimativa da produção antes do período da colheita, uma vez que poderá utilizar a previsão da safra para analisar as eventuais necessidades de transporte e armazenamento do produto, como também os possíveis lucros na sua comercialização (KHAIRUNNIZA-BEJO; MUSTAFFHA; ISMAIL, 2014). Em razão da facilidade do conhecimento quanto ao desempenho da cultura em determinado ambiente, na previsão e/ou exploração do potencial produtivo sob diversas condições, além da utilização para a gestão tecnológica e tomada de decisão, o desenvolvimento e a utilização de modelagem é uma metodologia significativa no setor agrícola (SOUSA et al., 2020).

Dentro desse contexto, como opção de modelagem e estimativa de variáveis de interesse, as Redes Neurais Artificiais (RNA) têm sido utilizadas com sucesso ao longo dos anos em culturas com alto potencial econômico, como tomate (MOVAGHARNEJAD; NIKZAD, 2007), batata (FORTIN et al., 2010), milho (FARJAM et al., 2014), soja (ALVES et al., 2018) e banana (GUIMARÃES et al., 2021).

Há aproximadamente oito décadas, as RNA vêm sendo estudadas e, atualmente, vem sendo empregadas em diversas áreas da ciência (SHAHID; RAPPON; BERTA, 2019). Foram inspiradas no Sistema Nervoso Biológico (Figura 2), uma vez que a percepção do cérebro humano realiza o processamento das informações de maneira mais ágil e distinta em comparação aos computadores digitais convencionais em virtude da sua complexidade de operação e ao fato de se processar em paralelo (HAYKIN, 2001).

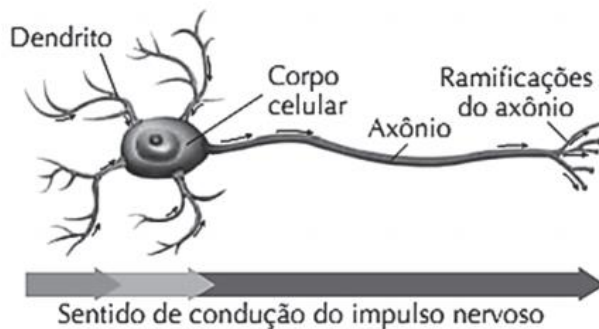


Figura 2. Representação de um neurônio biológico.

Fonte: <https://app.planejativo.com/q/17873/biologia-2/sistema-nervoso>

Em 1943, o primeiro modelo artificial de um neurônio biológico foi desenvolvido pelo neuroanatomista e psiquiatra Warren McCulloch e do matemático Walter Pitts, posteriormente aprimorado por Rosenblatt, que o nomeou de *Perceptron* (SILVA; SCHIMIDT, 2016). Após alguns anos de pesquisa, o modelo Perceptron Múltiplas Camadas (*Multilayer Perceptron* – MLP) (Figura 3) e o algoritmo *backpropagation* tornaram as Redes Neurais Artificiais uma metodologia amplamente utilizada em inúmeras áreas de estudo da ciência (BRAGA; CARVALHO; LUDERMIR, 2007). Portanto, seu desenvolvimento é marcado por muitas descobertas, conquistas e aperfeiçoamento de modelos que foram abandonados.

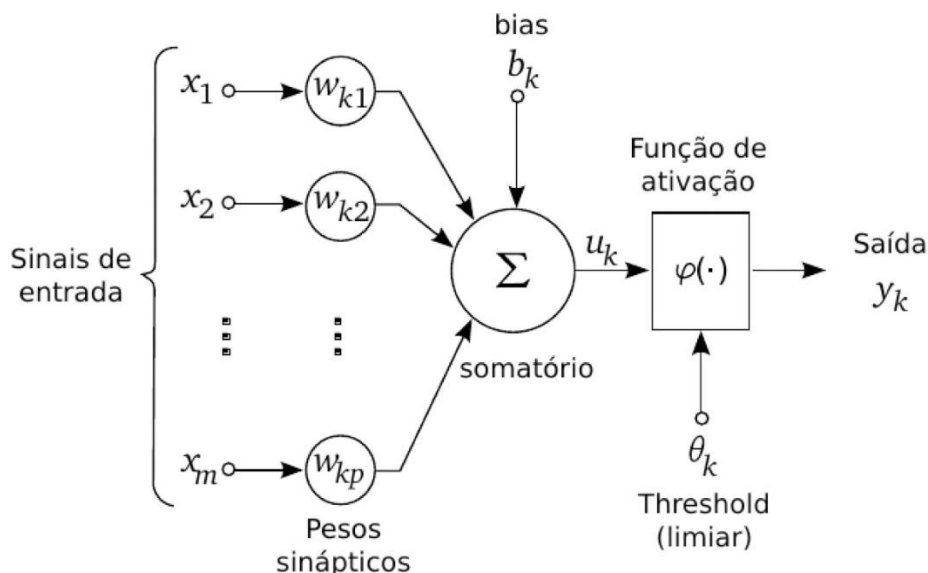


Figura 3. Modelo de neurônio artificial, segundo McCulloch-Pitts. Onde x_1, x_2, \dots, x_m são as entradas da rede; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos, ou pesos sinápticos, associados a cada entrada; b_k é o termo bias; u_k é a combinação linear dos sinais de entrada; $\varphi(\cdot)$ é a função de ativação e y_k é a saída do neurônio.

Fonte: SILVA; SCHIMIDT (2016).

Segundo Braga, Carvalho e Ludermir (2007), as RNA são métodos heurísticos que se assemelham às estruturas neurais biológicas (neurônios) e que têm capacidade computacional adquirida por meio de aprendizado e generalização.

O comportamento da rede é definido pelas conexões (pesos sinápticos) entre as suas camadas. Sendo assim, é possível treinar uma rede neural para desempenhar uma função particular ajustando-se os pesos das conexões entre as camadas (HAYKIN, 2001). O aprendizado de uma rede neural é um processo pelo qual os parâmetros são adaptados por meio de estímulos do ambiente onde a rede está inserida. Esse procedimento permite que a RNA adquira, através de uma melhora gradativa, uma boa capacidade de resposta para o maior número de situações possíveis. Ou seja, a RNA treinada torna-se capaz de generalizações (DINI; LORENZ; SOUZA, 2004).

As RNA caracterizam-se pela sua arquitetura e pelo ajustamento de seus pesos às conexões no decorrer do processo de aprendizado. A arquitetura de uma rede neural é constituída pelo número de camadas (camada única ou múltiplas camadas), pelas conexões entre camadas, pelo número de neurônios em cada camada, pelo tipo de conexão entre eles (*feedforward* ou *feedback*) e pelo algoritmo de aprendizado (IBM, 2020).

O número de camadas é o elemento essencial na determinação da capacidade de rede de resolver problemas. Geralmente, as camadas são classificadas em três categorias: camada de entrada (local em que os padrões são apresentados à rede); camadas intermediárias ou ocultas (designadas a efetuar grande parte do processamento dos dados e atribuir pesos mediante as conexões ponderadas, ou seja, são extratoras das características) e camada de saída (onde o resultado final é concluído e apresentado) (SINGH; PRAJNESHU, 2008; IBM, 2020).

A figura 4 representa uma arquitetura de rede neural na qual podem ser observadas a camada de entrada, a camada oculta e a camada de saída que retornam valores preditos para as variáveis de interesse, cuja resposta pode ser uni ou multivariado.

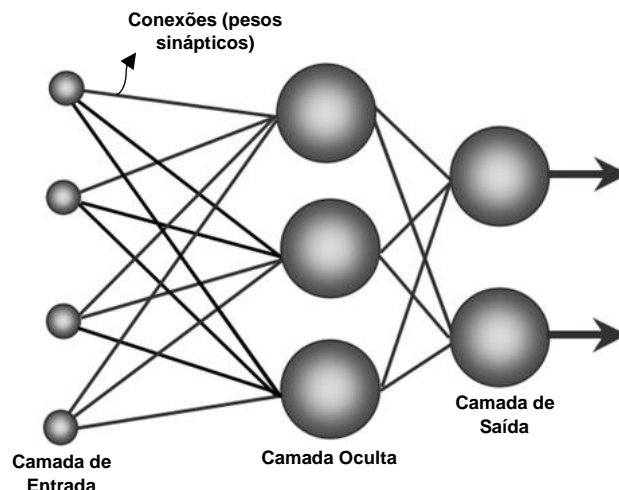


Figura 4. Rede Neural Artificial de múltiplas camadas.
Fonte: Adaptado de FURTADO (2019).

No decorrer do processo de treinamento das RNA, os pesos entre as conexões dos neurônios artificiais são parâmetros ajustáveis que variam à medida que o conjunto de treino é apresentado à rede. Ou seja, o processo de aprendizado em uma RNA resulta em sucessivos ajustes dos pesos sinápticos para que a saída da rede seja a mais próxima possível da resposta desejada. Logo, pode-se afirmar que é nos pesos sinápticos que se encontra todo o conhecimento adquirido pela rede. Além disso, o modelo neural também inclui um termo chamado de “bias”, simbolizado por bk , que tem o efeito do acréscimo ou decréscimo da função de ativação na entrada da rede, dependendo se é positiva ou negativa, respectivamente (HAYKIN, 2001).

2.4.1 Funções de ativação

Assim como nos neurônios biológicos, os neurônios artificiais também requerem estímulos que ultrapassem o valor limiar para que o impulso seja propagado. Desse modo, a porta de limiar funciona como um mecanismo para o disparo da resposta, isto é, compara o valor das somas ponderadas das entradas com um valor limite. Caso a soma exceda o limiar, a saída é ativada, caso não ocorra, ela permanece desativada (BRAGA; CARVALHO; LUDERMIR, 2007).

Portanto, as funções de ativação apresentam um papel importante de gerar o valor de saída do neurônio a partir das somas ponderadas recebidas pelo

neurônio, e devem ser escolhidas de acordo com o problema em estudo. Sendo que as mais utilizadas são: função limiar (degrau) (Figura 5a), função linear (purelin) (Figura 5b), função sigmoidal (logística) (Figura 5c) e função tangente hiperbólica (Figura 5d) (HAYKIN, 2001).

a. Função Limiar (Degrau)

A função limiar é aplicada em neurônios que tomam decisões binárias, limitando a saída do neurônio somente a dois valores. Então, os neurônios desenvolvidos com essa função, caso o valor de ativação v seja negativo, o valor da saída y será igual a 0, e 1 nos casos em que o valor de ativação for positivo (HAYKIN, 2001). A função limiar é expressada na seguinte equação:

$$\varphi(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{se } x < 0 \end{cases}$$

b. Função Linear (Purelin)

A função de ativação linear não limita a saída da RNA e é usada para armazenar entrada e saída de dados. Além disso, os neurônios que apresentam esta função atuam como aproximadores lineares (HAYKIN, 2001). Essa função é representada pela equação abaixo:

$$\varphi(x) = x$$

c. Função Sigmoidal (Logística)

De acordo com Haykin (2001), a função sigmoidal (logsig) é a função de ativação mais utilizada na construção das Redes Neurais Artificiais. Por definição, a função sigmoidal é monótona crescente com propriedades assintóticas e de suavidade. Ao contrário da função limiar, esta função pode atribuir todos os valores referentes ao intervalo de 0 e 1. A função sigmoidal é definida pela seguinte equação:

$$\varphi(x) = \frac{1}{1 + e^{-ax}}$$

Sendo que a é o parâmetro de inclinação da função sigmoidal. Onde a variação do valor do parâmetro a proporciona funções sigmoidais com diferentes inclinações.

d. Função Tangente Hiperbólica

De acordo com Haykin (2001), em certas circunstâncias é necessário que a função de ativação assuma valores entre 1 e -1, admitindo uma forma antissimétrica em relação à origem. Neste caso, utiliza-se um formato correspondente à logsig, denominada de função tangente hiperbólica (tansig). Essa função é representada pela equação abaixo:

$$\varphi(x) = \tanh\left(\frac{x}{2}\right) = \frac{1 - e^{-x}}{1 + e^x}$$

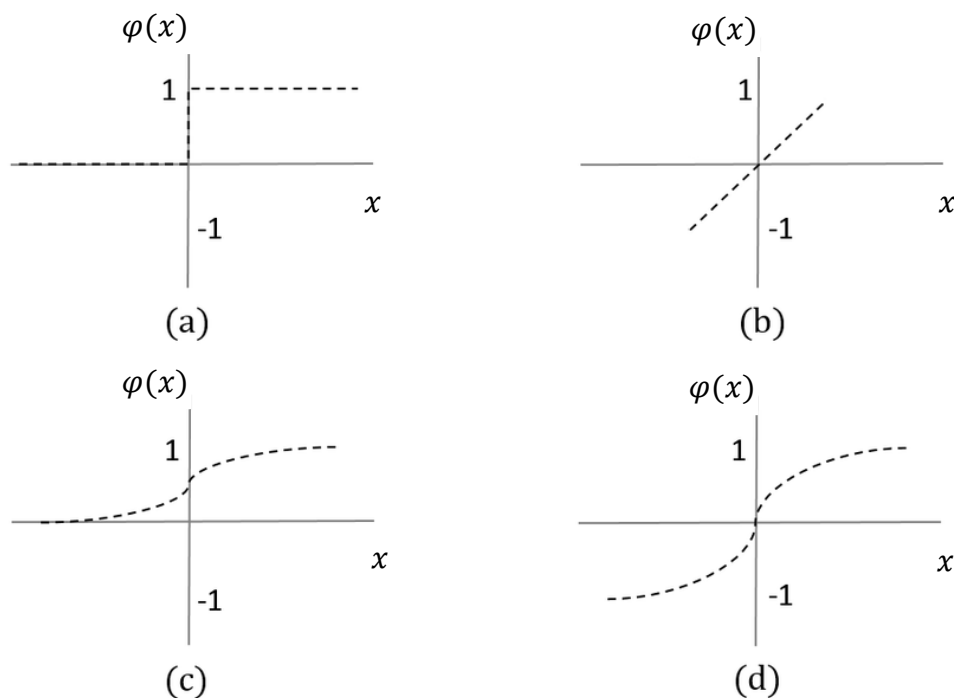


Figura 5. Representação gráfica das funções de ativação mais utilizadas: (a) função limiar; (b) função linear; (c) função sigmoidal e (d) função tangente hiperbólica.

Fonte: Adaptado de FIORIN et al. (2011).

2.4.2 Paradigmas de Aprendizagem

O processo de aprendizagem de uma rede neural consiste na etapa em que os parâmetros livres são adaptados por meio de um processo de estimulação fornecido pelo ambiente em que a rede está inserida, se tornando capaz de gerar uma solução generalizada para um determinado problema (HAYKIN, 2001).

O aprendizado é realizado por intermédio de processos iterativos de ajustes aplicados aos pesos sinápticos, o qual é denominado de treinamento (MIRANDA; FREITAS; FAGGION, 2009). Portanto, um treinamento adequado é a etapa mais importante do processamento dos dados, uma vez que os pesos atribuídos às conexões entre os neurônios detêm a aprendizagem e a capacidade de melhorar o desempenho (HAYKIN, 2001).

Segundo Bishop (1995), as RNA podem ser classificadas quanto à forma de aprendizagem, usando os seguintes paradigmas:

a. Aprendizado Supervisionado

A aprendizagem supervisionada (aprendizagem com professor), método mais comum para treinamento das RNA, é definida pelo emprego de um agente externo que informa à rede a resposta desejada para o padrão de entrada, como ilustrado na Figura 6. Assim, para cada entrada a rede gera uma resposta na saída, a qual é comparada com a resposta desejada, obtendo-se um erro referente à resposta atual. A partir disso, os pesos sinápticos são ajustados de forma a minimizar o erro, isto é, a cada resposta gerada pela rede, pequenos reparos são realizados nos pesos das conexões. O aprendizado supervisionado se aplica às classes de problemas em que se deseja mapear padrões de entrada e saída, como no caso de problemas de classificação e regressão (SILVA et al., 2017).

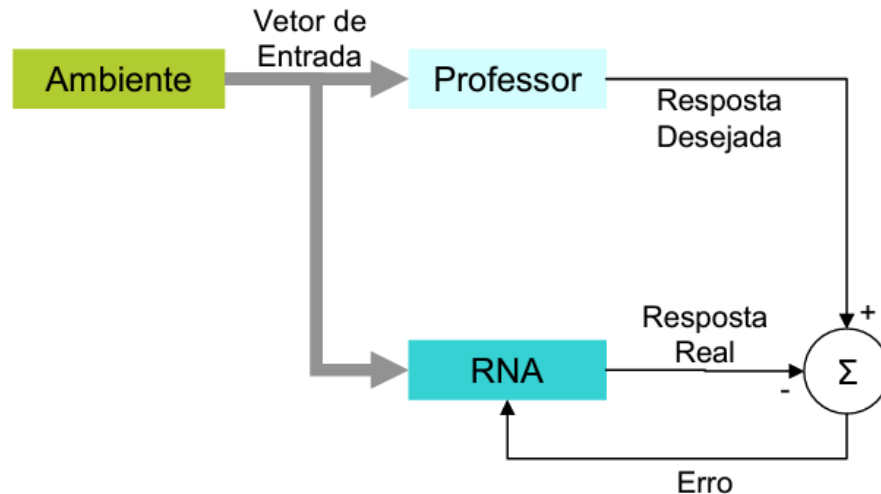


Figura 6. Mecanismo de aprendizado supervisionado.
Fonte: HAYKIN (2001).

b. Aprendizado Não Supervisionado

No processo de aprendizado não supervisionado, não existe um agente externo (professor/supervisor) para auxiliar no sistema de aprendizagem. Diante disso, o aprendizado não supervisionado não apresenta uma saída alvo, a própria rede deverá ser capaz de extrair as características relevantes dos impulsos, classificando-os em grupos pré-existentes. Assim, uma resposta é fornecida para cada entrada aplicada a rede, a fim de indicar a classe à qual pertence (Figura 7). Caso o padrão da entrada não corresponda às classes existentes, uma nova classe é originada. Dessa forma, o aprendizado não supervisionado se aplica às classes de problemas em que se deseja formar grupos de populações semelhantes, como no caso de problemas de *clusterização* (ou agrupamento) (SERRA; TAGLIAFERRI, 2019).



Figura 7. Mecanismo de aprendizado não supervisionado.
Fonte: HAYKIN (2001).

Um exemplo de rede neural não supervisionada é o Mapa Auto-Organizável de Kohonen (Figura 8), o qual é constituído por duas camadas que pode organizar os dados de entrada em um mapa topológico a partir de um início aleatório, realizando uma combinação da camada de entrada com uma camada competitiva de unidades processadoras. Assim, os padrões são classificados pelas unidades que ativam na camada de saída, mapeando as similaridades entre os padrões. Após o treinamento, a relação de similaridade entre os padrões é observada na camada competitiva (HAYKIN, 2009).

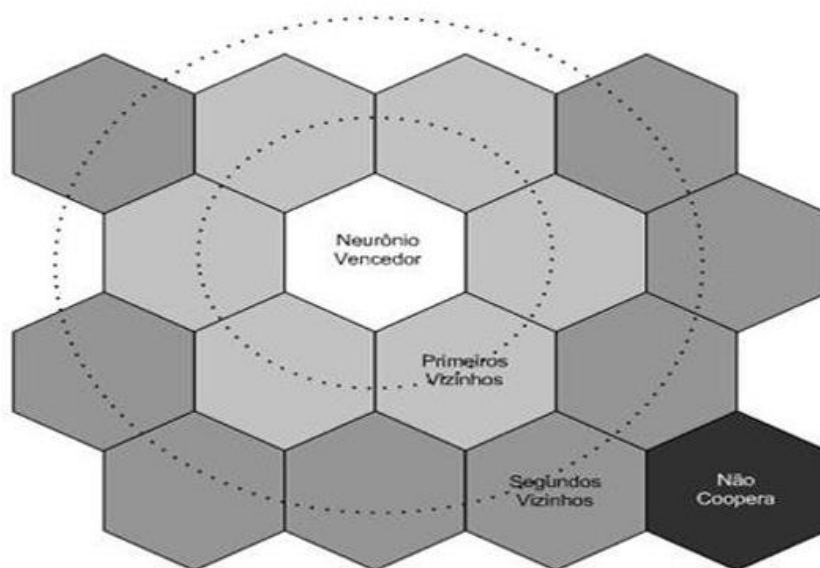


Figura 8. Representação de um Mapa Auto-Organizável de Kohonen.
Fonte: ICMC (2009).

De acordo com Haykin (2009), a formação do Mapa Auto-Organizável de Kohonen pode ser compreendida em três processos:

- I. **Competição:** Para cada padrão de entrada, os neurônios da rede calculam seus respectivos valores mediante a função discriminante, a qual fornece a base para a competição entre os neurônios. O neurônio com o maior valor de função discriminante é o vencedor da competição.
- II. **Cooperação:** O neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo, desta forma, as bases para a cooperação entre os neurônios vizinhos.

- III. Adaptação Sináptica: Os neurônios excitados aumentam seus valores individuais da função discriminante em relação ao padrão de entrada mediante os ajustes adequados aplicados a seus pesos sinápticos. Conseqüentemente, a grade vai se tornando auto-organizável.

2.4.3 Vantagens e Dificuldades da implementação das Redes Neurais Artificiais

Para que uma RNA consiga obter um bom desempenho é necessário que faça uma escolha adequada dos seus atributos de acordo com o problema a ser solucionado (MARTINS et al., 2016). De acordo com Basheer e Hajmeer (2000), alguns parâmetros são determinados por tentativa e erro, isto é, são atribuídos vários valores distintos aos parâmetros e os resultados obtidos são analisados, a melhor configuração é escolhida.

Dentre as dificuldades encontradas no treinamento (ou aprendizagem) de uma RNA estão: número de camadas da rede treinada, a quantidade de neurônios presentes nas camadas ocultas, e as funções de ativação que estarão presentes tanto nas camadas ocultas quanto nas camadas de saída (MARTINS et al., 2016). Além da determinação do número ideal de ciclos (ou épocas) de treinamento da rede (BASHEER; HAJMEER, 2000). Uma vez que se um número muito grande de ciclos de treinamento for aplicado, a rede entra em um processo de memorização dos padrões a ela apresentados, denominado de *overfitting*, perdendo assim a capacidade de generalização e extrapolando suas conclusões para os dados não treinados. Assim, o supertreinamento é identificado quando o erro, obtido pela validação, começa a aumentar depois de ter diminuído. Para evitar o *overfitting*, a aprendizagem deve ser interrompida quando os erros não variam mais, indicando que a rede está treinada (JABBAR; KHAN, 2015). Esse fenômeno pode ser observado na Figura 9.

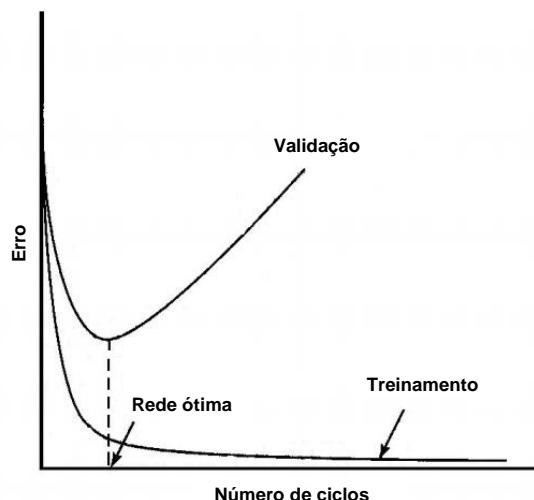


Figura 9. Representação gráfica do número de ciclos para o treinamento da rede.

Fonte: Adaptado de BASHEER; HAJMEER (2000).

Enquanto que se um número muito pequeno for aplicado, a rede torna-se incapaz de representar os dados, visto que a rede pode não encontrar a solução mais adequada para o problema em questão (*underfitting*) (JABBAR; KHAN, 2015).

Dentre as vantagens da utilização das RNA, ressaltam-se os seguintes fundamentos: a sua estrutura não linear, visto que as redes neurais artificiais são amplamente aplicadas em estudos em que as relações entre as variáveis são desconhecidas ou muito complexas (GAGO et al., 2010), apresentando assim a capacidade de não requerer informação detalhada sobre os processos físicos do sistema a ser modelado (SUDHEER; GOSAIN; RAMASASTRI, 2003). Além disso, as RNA possibilitam a modelagem com variáveis qualitativas (ou categóricas) de forma eficiente (MARTINS et al., 2016). Devido a esses atributos, o emprego das redes neurais tem se destacado promissora em virtude da possibilidade de um desempenho superior aos modelos convencionais utilizados na solução de problemas (BRAGA; CARVALHO; LUDERMIR, 2007).

2.4.4 Aplicação das Redes Neurais Artificiais no melhoramento

Sousa et al. (2020) compararam os modelos de Regressão Linear Múltipla (MLR) e Rede Neural Artificial (RNA) na predição da produtividade de grãos de

dez variedades tradicionais de feijão-fava. Além disso, avaliaram a adaptabilidade e a estabilidade mediante ao método de Lin e Binns para identificação da variedade com melhor desempenho. Os parâmetros utilizados para comparar os modelos foram: Correlação de Pearson e Spearman, Raiz Quadrada do Erro Médio (RQEM), e Erro Absoluto Médio (EAM). Assim, concluíram que a RNA se mostrou mais adequada para predição da produtividade de grãos.

Santos et al. (2017) testaram a aplicabilidade das Redes Neurais Artificiais na estimação da altura em testes clonais e de progênies de eucalipto. Tanto para a seleção do teste clonal quanto de progênies, foram utilizadas duas sub-amostras para a definição dos dados para o treinamento das redes, sendo que a sub-amostra 1 constituía de dados da primeira árvore de cada bloco enquanto que a sub-amostra 2 constituía de dados de uma árvore selecionada aleatoriamente dentro de cada bloco. Assim, os resultados demonstraram que as RNA são eficientes na estimativa da altura das árvores dos diferentes genótipos dos testes clonais e de progênies, permitindo uma redução do tempo de trabalho na medição de parcelas experimentais e, conseqüentemente, uma redução nos custos da pesquisa, sem perder a precisão da medição.

Peixoto, Bhering e Cruz (2015) avaliaram a eficiência de RNA na predição do valor genético em experimentos com estrutura de blocos ao acaso. Para analisar a precisão das RNA, houve uma avaliação comparando a correlação do valor da rede com o valor genético bem como do valor fenotípico com o valor genético. Assim, os resultados demonstraram que as redes neurais foram eficientes em prever o valor genético com um ganho de 0,64 a 10,03% em relação ao valor fenotípico, independentemente do tamanho da população simulada, herdabilidade ou coeficiente de variação. Portanto, concluíram que a rede neural artificial é uma técnica promissora para prever o valor genético em experimentos balanceados em blocos aleatórios para características quantitativas (fortemente influenciadas pelo ambiente).

Barbosa et al. (2011) analisaram 37 acessos de mamão (*Carica papaya* L.) e oito características quantitativas com o objetivo de avaliar a viabilidade da RNA, modelo baseado em Kohonen (1982), como técnica alternativa para

análises da diversidade genética. Após a obtenção dos resultados, houve uma comparação entre as Redes Neurais Artificiais de Kohonen e a Análise Discriminante de Anderson, onde pôde-se observar que, de acordo com a análise discriminante de Anderson, 91,90% dos acessos foram classificados corretamente nos grupos previamente definidos pela RNA. Então, concluíram que a técnica da RNA de Kohonen se demonstrou viável na classificação dos acessos, visto que se constatou a presença significativa de diversidade genética entre os acessos avaliados.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABRAMO, M. A. **Taioba, cará e inhame: o grande potencial inexplorado**. São Paulo: Editora Ícone, 1990. 80 p.
- AKORODA, M. O. Floral biology in relation to hand pollination of white yam. **Euphytica**, v. 32, n. 3, p. 831-838, 1983.
- ALEXANDER, J.; COURSEY, D. G. The origins of yam cultivation. In: PETER, U. J.; DIMBLEBY, G. W (org.). **The domestication and exploitation of plants and animals**. Londres, Cerald Duckworth, 1969. 405 p.
- ALVES, G. R.; TEIXEIRA, I. R.; MELO, F. R.; SOUZA, R. T. G.; SILVA, A. G. Estimating soybean yields with artificial neural networks. **Acta Scientiarum. Agronomy**, v. 40, 2018.
- ANDRES, C.; ADEOLUWA, O. O.; BHULLAR, G. S. Yam (*Dioscorea* spp.). **Encyclopedia of Applied Plant Sciences**, v. 3, p. 435–441, 2017.
- ANOKYE, M.; TETTEH, J. P.; OTOO, E. Morphological characterization of some water yam (*Dioscorea alata* L.) germplasm in Ghana. **Journal of Agricultural Science and Technology**, v. 4, p. 518-532, 2014.
- ASIEDU, R.; WANYERA, N. M.; NG, S. Y. C.; NG, N. Q. Yams. In: FUCCILLO, D.; SEARS, L.; STAPLETON, P. (org.). **Biodiversity in trust: Conservation and use of plant genetic resources in CGIAR centers**. Cambridge University Press, Cambridge, UK, 1997. 57-66 p.
- ATIENO, V.; GATHERI, G. W.; KAMAU, J. W.; MUTHINI, M. Morphological and molecular characterization of cultivated yam (*Dioscorea* species) in selected counties in Kenya. **African Journal of Plant Science**, v. 14, n. 7, p. 270-279, 2020.
- BARBOSA, C. D.; VIANA, A. P.; QUINTAL; S. S. R.; PEREIRA, M. G. Artificial neural network analysis of genetic diversity in *Carica papaya* L. **Crop Breeding and Applied Biotechnology**, v. 11, p. 224-231, 2011.
- BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. **Journal of microbiological methods**, v. 43, n. 1, p. 3-31, 2000.
- BATISTA, N. N.; RAMOS, C. L.; PIRES, J. F.; MOREIRA, S. I.; ALVES, E.; DIAS, D. R.; SCHWAN, R. F. Nondairy ice cream based on fermented yam (*Dioscorea* sp.). **Food science & nutrition**, v. 7, n. 5, p. 1899-1907, 2019.
- BATISTA, N. N.; RAMOS, C. L.; VILELA, L. F.; DIAS, D. R.; SCHWAN, R. F. Fermentation of yam (*Dioscorea* spp. L.) by indigenous phytase-producing lactic acid bacteria strains. **Brazilian Journal of Microbiology**, v. 50, n. 2, p. 507-514, 2019.
- BISHOP, C. M. Neural Networks for Pattern Recognition. Birmingham, UK: Clarendon Press, Oxford, 1995. 498 p. Disponível em: <<http://people.sabanciuniv.edu/berrin/cs512/lectures/Book-Bishop->

[Neural%20Networks%20for%20Pattern%20Recognition.pdf](#)>. Acesso em: 12 Jul. 2021.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. LTC Editora, 2ª edição, 2007. 226 p.

BRASIL. Alimentos Regionais Brasileiros. Brasília – DF: Ministério da Saúde, 2015. 484 p. Disponível em: <https://bvsms.saude.gov.br/bvs/publicacoes/alimentos_regionais_brasileiros_2ed.pdf>. Acesso em: 17 Jun. 2021.

COURSEY, D. G. The origin and domestication of yams in Africa. In: HARLAN, J. R.; WET, J. M. J.; MOUTON, A. B. L. S. (org.). **Origin of African plant domestication**. Amsterdam: Mouton Hague, 1976. 383-408 p.

COURSEY, D. G. **Yams. An account of the nature, origins, cultivation and utilization of the useful members of dioscoreaceae**. Londres: Longmans, Green and Co. Ltd, 1967. 78-91 p.

COURSEY, D. G. Yams *Dioscorea* spp. (Dioscoreaceae). In: SIMMONDS, N. W. (org.). **Evolution of crop plants**. Longman Group, 1976. 70-74 p.

COUTO, R. S.; MARTINS, A. C.; BOLSON, M.; LOPES, R. C.; SMIDT, E. C.; BRAGA, J. M. A. Time calibrated tree of *Dioscorea* (Dioscoreaceae) indicates four origins of yams in the Neotropics since the Eocene. **Botanical Journal of the Linnean Society**, v. 188, n. 2, p. 144-160, 2018.

DINI, P.; LORENZ, P.; SOUZA, J. N. **Service Assurance with Partial and Intermittent Resources**. New York, USA: Springer-Verlag Berlin Heidelberg New York, 2004. 311 p.

EPPING, J.; LAIBACH, N. An underutilized orphan tuber crop—Chinese yam: a review. **Planta**, v. 252, n. 4, p. 1-19, 2020.

FARJAM, A.; OMID, M.; AKRAM, A.; NIARI, Z. F. A neural network based modeling and sensitivity analysis of energy inputs for predicting seed and grain corn yields. **Journal of Agricultural Science and Technology**, v. 16, n. 4, p. 767-778, 2014.

FIORIN, D. V.; MARTINS, F. R.; SCHUCH, N. J.; PEREIRA, E. B. Aplicações de redes neurais e previsões de disponibilidade de recursos energéticos solares. **Revista Brasileira de Ensino de Física**, v. 33, p. 01-20, 2011.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). Participation of continents in yam production in 2019. FAOSTAT Database Gateway – FAO. Disponível em: <<http://www.fao.org/faostat/en/#data/QC>>. Acesso em: 21 abr. 2021.

FORTIN, J. G.; ANCTIL, F.; PARENT, L.; BOLINDER, M. A. A neural network experiment on the site-specific simulation of potato tuber growth in Eastern Canada. **Computers and Electronics in Agriculture**, v. 73, n. 2, p. 126-132, 2010.

FURTADO, M. I. V. **Redes Neurais Artificiais: Uma abordagem para sala de aula**. Ponta Grossa (PR): Atena Editora, 2019. 99 p.

GAGO, J.; MATÍNEZ-NÚÑEZ, L.; LANDÍN, M.; GALLEGO, P. P. Artificial neural networks as an alternative to the traditional statistical methodology in plant research. **Journal of plant physiology**, v. 167, n. 1, p. 23-27, 2010.

GUIMARÃES, B. V. C.; DONATO, S. L. R.; ASPIAZÚ, I.; AZEVEDO, A. M. Yield prediction of 'Prata Anã' and 'BRS Platina' banana plants by artificial neural networks. **Pesquisa Agropecuária Tropical**, v. 51, 2021.

HAYKIN, S. **Neural Networks and Learning Machines**. Hamilton, Ontario, Canadá: Pearson Prentice Hall, 2009. 938 p.

HAYKIN, S. **Redes Neurais: Princípios e Práticas**. Porto Alegre: Bookman, 2001. 900 p.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). Levantamento Sistemático da Produção Agrícola (LSPA). Disponível em: <<https://sidra.ibge.gov.br/tabela/6619>>. Acesso em: 26 abr. 2021.

INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO (ICMC). Redes Neurais Artificiais. 2009. Disponível em: <<https://sites.icmc.usp.br/andre/research/neural/>>. Acesso em: 17 Nov. 2021.

INTERNATIONAL PLANT GENETIC RESOURCES INSTITUTE (IPGRI)/INTERNATIONAL INSTITUTE OF TROPICAL AGRICULTURE (IITA). Descriptors for yam (*Dioscorea* spp.). Roma, Itália: IPGRI/IITA, 1997. 61 p. Disponível em: <https://www.biodiversityinternational.org/fileadmin/_migrated/uploads/tx_news/Descriptors_for_Yam__Dioscorea_spp.__310.pdf>. Acesso em: 15 Jun. 2021.

INTERNATION BUSINESS MACHINES CORPORATION (IBM). Neural Networks. Disponível em: <<https://www.ibm.com/cloud/learn/neural-networks>>. Acesso em: 27 Jun. 2021.

JABBAR, H.; KHAN, R. Z. Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). **Computer Science, Communication and Instrumentation Devices**, p. 163-172, 2015.

KHAIRUNNIZA-BEJO, S.; MUSTAFFHA, S.; ISMAIL, W. I. W. Application of artificial neural network in predicting crop yield: A review. **Journal of Food Science and Engineering**, v. 4, n. 1, p. 1, 2014.

KULASINGHEA, W. M. A. A.; RANAWEERA, K. K. T. N. Physical, chemical and biological aspects of *Dioscorea* yams and potential value additions. **Journal of Agriculture and Value Addition**, v. 2, p. 43–59, 2019.

MARTINS, E. R.; BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para estimação da altura total de árvores de eucalipto. **Revista brasileira de ciências agrárias**, v. 11, n. 2, p. 117-123, 2016.

MENDES, L. N.; SILVA, J. A.; FAVERO, L. A. Panorama da produção e comercialização do inhame no mundo e no Brasil e sua importância para o mercado pernambucano: uma análise das cinco forças competitivas. In:

CONGRESSO ONLINE DE ADMINISTRAÇÃO, 10., 2013. **Anais...** 2013, 1-15 p.

MESQUITA, A. S. Inhame e taro: Cenários dos mercados internacional, brasileiro e baiano. **Bahia Agrícola**, v. 5, n. 2, p. 54-64, 2002.

MIGNOUNA, H. D.; ABANG, M. M.; ASIEDU, R. Advances in yam (*Dioscorea* spp.) genetics and genomics. **Proceedings of the 13th ISTRC Symposium**, Arusha, p. 72-81, 2007.

MIRANDA, F. A.; FREITAS, S. R. C.; FAGGION, P. L. Integração e interpolação de dados de anomalias ar livre utilizando-se a técnica de RNA e krigagem. **Boletim de Ciências Geodésicas**, v. 15, n. 3, p. 428-443, 2009.

MONTALDO, A. **Cultivo de raíces y tubérculos tropicales**. San José, Costa Rica: Instituto Interamericano de Cooperación para la Agricultura, 1991. 408 p.

MONDO, J. M.; EDEMODOU, A.; ADEBOLA, P.; ASIEDU, R.; AKORODA, M. O.; ASFAW, A. Floral Biology and Pollination Efficiency in Yam (*Dioscorea* spp.). **Agriculture**, v. 10, n. 11, p. 560, 2020.

MOVAGHARNEJAD, K.; NIKZAD, M. Modeling of tomato drying using artificial neural network. **Computers and electronics in agriculture**, v. 59, n. 1-2, p. 78-85, 2007.

NORONHA, M. A. **Principais Doenças do Inhame (*Dioscorea cayennensis*) nos Tabuleiros Costeiros do Nordeste**. Aracaju – SE: EMBRAPA Tabuleiros Costeiros, 2014. 41 p.

OBIDIEGWU, J. E.; AKPABIO, E. M. The geography of yam cultivation in southern Nigeria: Exploring its social meanings and cultural functions. **Journal of Ethnic Foods**, v. 4, n. 1, p. 28-35, 2017.

OBIDIEGWU, J. E.; KOLESNIKOVA-ALLEN, M.; ENE-OBONG, E. E.; MUONEKE, C. O.; ASIEDU, R. SSR markers reveal diversity in Guinea yam (*Dioscorea cayenensis/D. rotundata*) core set. **African Journal of Biotechnology**, v. 8, n. 12, 2009.

ODU, N. N.; ELENWO, M.; MADUKA, N. Microbiological quality of packaged and exposed cassava, yam and plantain flour sold in markets and supermarkets in Port Harcourt metropololis, Nigeria. **American Journal of Microbiological Research**, v. 7, n. 2, p. 57-62, 2019.

OMOHIMI, C.; PICCIRILLO, C.; FERRARO, V.; RORIZ, M. C.; OMEMU, M. A.; SANTOS, S. M.; RESSUREIÇÃO, S.; ABAYOMI, L.; ADEBOWALE, A.; VASCONCELOS, M. W.; OBADINA, O.; SANNI, L.; PINTADO, M. M. E. Safety of yam-derived (*Dioscorea rotundata*) foodstuffs—chips, flakes and flour: effect of processing and post-processing conditions. **Foods**, v. 8, n. 1, p. 12, 2019.

PADHAN, B.; PANDA, D. Potential of neglected and underutilized yams (*Dioscorea* spp.) for improving nutritional security and health benefits. **Frontiers in pharmacology**, v. 11, p. 496, 2020.

PEDRALLI, G. Dioscoreaceae. In: RIBEIRO, J. E.; HOPKINS, M.; VICENTINI, A. (org.). **Flora da Reserva Ducke: Guia de Identificação das plantas**

- vasculares de uma floresta de terra firme da Amazônia.** Manaus, AM: INPA/DFID, 1999. 723-724 p.
- PEIXOTO, L. A.; BHERING, L. L.; CRUZ, C. D. Artificial neural networks reveal efficiency in genetic value prediction. **Genetics and Molecular Research**, v. 14, n. 2, p. 6796-6807, 2015.
- RAMAN, V.; GALAL, A. M.; AVULA, B.; SAGI, S.; SMILLIE, T. J.; KHAN, I. A. Application of anatomy and HPTLC in characterizing species of *Dioscorea* (Dioscoreaceae). **Journal of natural medicines**, v. 68, n. 4, p. 686-698, 2014.
- RAMSER, J.; WEISING, K.; TERAUCHI, R.; KAHL, G.; LOPEZ-PERALTA, C.; TERHALLE, W. Molecular marker based taxonomy and phylogeny of Guinea yam (*Dioscorea rotundata*–*D. cayenensis*). **Genome**, v. 40, n. 6, p. 903-915, 1997.
- SANTOS, A. C. A.; ALMEIDA, F. M.; SOUZA, R. B.; CHAVES, R.; PAIVA, H. N.; BINOT, D. H. B.; LEITE, H. G.; FARIAS, A. A. Estimation of eucalyptus tree height in clonal and progeny tests using artificial neural networks. **Revista Árvore**, v. 41, n. 6, p. e410602, 2017.
- SANTOS, E. S. Manejo Sustentável da cultura do inhame (*Dioscorea* sp.) no Nordeste do Brasil. In: SIMPÓSIO NACIONAL SOBRE AS CULTURAS DE INHAME E TARO, 2., 2002, João Pessoa, PB. **Anais...** João Pessoa, PB: EMEPAPB, v. 1, 2002, 181-195 p.
- SCARCELLI, N.; CUBRY, P.; AKAKPO, R.; THUILLET, A.; OBIDIEGWU, J.; BACO, M. N.; OTOO, E.; SONKÉ, B.; DANSI, A.; DJEDATIN, G.; MARIAC, C.; COUDERC, M.; CAUSSE, S.; ALIX, K.; CHAIR, H.; FRANÇOIS, O.; VIGOUROUX, Y. Yam genomics supports West Africa as a major cradle of crop domestication. **Science advances**, v. 5, n. 5, p. eaaw1947, 2019.
- SHEIKH, N.; KUMAR, Y. Morphological Characterization of Meghalayan *Dioscorea* spp. (yam), North East India. **Journal of Agricultural Science And Technology**, v. 19, n. 2, 2017.
- SILVA, A. M. L.; NETO, M. L. F.; FILHO, F. C. I.; MOURA, R. S.; MACHADO, V. P. Descoberta de conhecimento através de métodos de aprendizagem de máquina supervisionados aplicados ao SIGAA/UFPI. **Revista de Sistemas e Computação-RSC**, v. 7, n. 1, 2017.
- SILVA, M. J.; JÚNIOR, A. C. S.; CRUZ, C. D.; NASCIMENTO, M.; OLIVEIRA, M. S.; SCHAFFERT, R. E.; PARRELLA, R. A. C. Computational intelligence for studies on genetic diversity between genotypes of biomass sorghum. **Pesquisa Agropecuária Brasileira**, v. 55, 2020.
- SILVA, S. R.; SCHIMIDT, F. Redução de Variáveis de Entrada de Redes Neurais Artificiais a Partir de Dados de Análise de Componentes Principais na Modelagem de Oxigênio Dissolvido. **Química Nova**, v. 39, n. 3, p. 273-278, 2016.
- SIQUEIRA, M. V. B. M. Inhame (*Dioscorea* spp.): Uma cultura ainda negligenciada. **Horticultura Brasileira**, v. 27, n. 2, 2009.

SIQUEIRA, M. V. B. M.; NASCIMENTO, W. F.; SILVA, L. R. G.; FERREIRA, A. B.; SILVA, E. F.; MING, L. C.; VEASEY, E. A. Distribution, management and diversity of yam local varieties in Brazil: a study on *Dioscorea alata* L. **Brazilian Journal of Biology**, v. 74, n. 1, p. 52-61, 2014.

SIQUEIRA, M. V. B. M. Yam: a neglected and underutilized crop in Brazil. **Horticultura Brasileira**, v. 29, n. 1, p. 16-20, 2011.

SINGH, R. K.; PRAJNESHU. Artificial neural network methodology for modelling and forecasting maize crop yield. **Agricultural Economics Research Review**, v. 21, n. 347-2016-16813, p. 5-10, 2008.

SERRA, A.; TAGLIAFERRI, R. Unsupervised Learning: Clustering. **Encyclopedia of Bioinformatics and Computational Biology**, v. 1, p. 350-357, 2019.

SHAHID, N.; RAPPON, T.; BERTA, W. Applications of artificial neural networks in health care organizational decision-making: A scoping review. **PloS one**, v. 14, n. 2, p. e0212356, 2019.

SOUSA, A. M. C. B.; SILVA, V. B.; LOPES, A. C. A.; GOMES, R. L. F.; CARVALHO, L. C. B. Prediction of grain yield, adaptability, and stability in landrace varieties of lima bean (*Phaseolus lunatus* L.). **Crop Breeding and Applied Biotechnology**, v. 20, n. 1, 2020.

SOUSA, E. C.; RAIZADA, M. N. Contributions of African crops to American culture and beyond: The Slave Trade and other journeys of resilient peoples and crops. **Frontiers in Sustainable Food Systems**, v. 4, p. 272, 2020.

SUDHEER, K. P.; GOSAIN, A. K.; RAMASASTRI, K. S. Estimating actual evapotranspiration from limited climatic data using neural computing technique. **Journal of irrigation and drainage engineering**, v. 129, n. 3, p. 214-218, 2003.

TRINDADE, T.; SOARES, L. S.; FURTADO, M. C.; CASTRO, A. A.; CARNELOSSI, M. A. G. Composição centesimal de inhame (*Dioscorea* sp.) *in natura* e minimamente processado. **Scientia Plena**, v. 7, n. 6, 2011.

UNITED STATES DEPARTMENT OF AGRICULTURE (USDA). Plant Guide. Foreign Agricultural Service, 2016. Disponível em: <<https://www.nass.usda.gov/AgCensus/>>. Acesso em: 26 abr. 2021.

ZOUNDJIHEKPON, J.; HAMON, P.; NOIROT, M.; TIO-TOURE, B.; HAMON, S. Flowering synchronisation between male and female West African cultivated yams (*Dioscorea cayenensis-rotundata* complex). **Euphytica**, v. 95, n. 3, p. 371-375, 1997.

CAPÍTULO I

REDES NEURAIS ARTIFICIAIS PARA A PREDIÇÃO DO PESO DO TUBÉRCULO DE INHAME (*Dioscorea rotundata*)

REDES NEURAIIS ARTIFICIAIS PARA A PREDIÇÃO DO PESO DO TUBÉRCULO DE INHAME (*Dioscorea rotundata*)

Autor: Ingrid Giovanna Vieira Santos

Orientador: Prof. Dr. Ricardo Franco Cunha Moreira

Coorientador: Prof. Dr. Liniker Fernandes da Silva

RESUMO: O objetivo deste estudo foi avaliar a predição do peso do tubérculo de inhame (*Dioscorea rotundata*) mediante a análise de Redes Neurais Artificiais. Para isso, foram utilizados os dados da caracterização agrônômica de 120 genótipos, os quais foram divididos aleatoriamente em duas partes, sendo 75% para o treinamento e 25% para a validação da RNA. Após a separação dos dados, foram calculadas as estatísticas descritivas das variáveis quantitativas: valores mínimo e máximo, média, mediana, desvio padrão e coeficiente de variação. E, para obter as RNA do tipo *Multilayer Perceptron* (MLP), utilizou-se o pacote *h2o* do *software* RStudio versão 1.4.1103. Foram treinadas um total de 1000 RNA, e destas, foram selecionadas as 5 melhores para serem apresentadas. Desse modo, avaliou-se a qualidade das RNA nas etapas de treino e validação mediante as seguintes estatísticas: Raiz do Erro Quadrático Médio (REQM), Erro Absoluto Médio (EAM) e análise gráfica de resíduos. Assim, os resultados do estudo evidenciaram a RNA como uma ferramenta satisfatória na predição do peso do tubérculo do inhame.

PALAVRAS-CHAVE: Inteligência artificial; Modelagem; *Multilayer Perceptron*; Agricultura; Programas de melhoramento.

ARTIFICIAL NEURAL NETWORKS FOR THE PREDICTION OF YAM TUBER WEIGHT (*Dioscorea rotundata*)

Author: Ingrid Giovanna Vieira Santos

Advisor: Prof. Dr. Ricardo Franco Cunha Moreira

Co-Advisor: Prof. Dr. Liniker Fernandes da Silva

ABSTRACT: The aim of this study was to evaluate the prediction of tuber weight of yam (*Dioscorea rotundata*) through the analysis of Artificial Neural Networks. For this, data from the agronomic characterization of 120 genotypes were used, which were randomly divided into two parts, 75% for training and 25% for ANN validation. After data separation, descriptive statistics of quantitative variables were calculated: minimum and maximum values, mean, median, standard deviation and coefficient of variation. And, to obtain the ANN of the Multilayer Perceptron (MLP) type, the h2o package of the RStudio software version 1.4.1103 was used. A total of 1000 ANN were trained, and from these, the 5 best were selected to be presented. Thus, the quality of the ANNs in the training and validation stages was evaluated using the following statistics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and graphic analysis of residues. Thus, the results of the study showed ANN as a satisfactory tool in predicting yam tuber weight.

KEYWORDS: Artificial intelligence; Modeling; Multilayer Perceptron; Agriculture; Improvement programs.

1. INTRODUÇÃO

O inhame é uma planta pertencente à família Dioscoreaceae e o gênero *Dioscorea*, o qual está distribuído nas regiões tropicais, subtropicais e temperadas de todo mundo, e constitui uma excelente fonte nutricional (EPPING; LAIBACH, 2020). Além do valor alimentar, o inhame tem se destacado como insumo na fabricação de cosméticos e nas indústrias farmacêuticas, sendo utilizado na medicina tradicional e contemporânea, apresentando relevância social, cultural e econômica nos países em que é amplamente cultivado, visto que muitas pessoas dependem para alimentação e subsistência (ASIEDU; SARTIE, 2010).

O desenvolvimento e a utilização de modelagem nas atividades agrícolas tornaram-se uma ferramenta fundamental para auxiliar a gestão tecnológica e a tomada de decisão, bem como a previsão e/ou exploração do potencial produtivo da cultura (SOUSA et al., 2020). Dentro desse contexto, como opção de modelagem e estimativa de variáveis de interesse, as Redes Neurais Artificiais (RNA) têm sido utilizadas com sucesso ao longo dos anos em culturas com alto potencial econômico, como tomate (MOVAGHARNEJAD; NIKZAD, 2007), batata (FORTIN et al., 2010), milho (FARJAM et al., 2014), soja (ALVES et al., 2018) e banana (GUIMARÃES et al., 2021).

As RNA são métodos heurísticos inspirados na arquitetura neural do cérebro humano, apresentando capacidade de aprendizagem mediante a experiência, ou seja, é capaz de reconhecer padrões, e de generalizar a informação aprendida, gerando um modelo não linear (SOUSA et al., 2020). Uma RNA consiste em múltiplas unidades de processamento simples, denominados neurônios artificiais, que estão conectados entre si e organizados em camadas de entrada, intermediária ou oculta e de saída, constituindo um sistema computacional paralelo para executar uma determinada tarefa (PFEIFER; DAMIAN; FUCHSLIN, 2010).

As RNA caracterizam-se pela sua arquitetura e pelo ajustamento de seus pesos às conexões no decorrer do processo de aprendizado. A arquitetura de uma rede neural é constituída pelo número de camadas (camada única ou

múltiplas camadas), pelas conexões entre camadas, pelo número de neurônios em cada camada, pelo tipo de conexão entre eles (*feedforward* ou *feedback*) e pelo algoritmo de aprendizado (IBM, 2020).

No decorrer do processo de treinamento das RNA, os pesos entre as conexões dos neurônios artificiais são parâmetros ajustáveis que variam à medida que o conjunto de treino é apresentado à rede. Ou seja, o processo de aprendizado em uma RNA resulta em sucessivos ajustes dos pesos sinápticos para que a saída da rede seja a mais próxima possível da resposta desejada. Logo, pode-se afirmar que é nos pesos sinápticos que se encontra todo o conhecimento adquirido pela rede (HAYKIN, 2001).

Segundo Murase (2000), o emprego desse método de inteligência computacional tem promovido respostas significativas nos sistemas de produção agrícola. O êxito da aplicação das RNA está diretamente relacionado com a sua alta versatilidade, uma vez que é utilizada em diversas áreas da ciência, e isso faz delas uma tecnologia promissora para o desenvolvimento da agricultura.

Porém, há uma escassez de trabalhos com a aplicabilidade da RNA como ferramenta preditiva de expressão fenotípica de características agronômicas em inhame. Portanto, o presente trabalho teve como objetivo avaliar a predição do peso do tubérculo do inhame (*Dioscorea rotundata*) mediante a análise de Redes Neurais Artificiais.

2. MATERIAL E MÉTODOS

2.1 Área de Estudo

Um total de 120 genótipos foram coletados em 18 propriedades rurais de povoados pertencentes aos municípios de Cruz das Almas (39° 06' 07" W, 12° 40' 12" S), São Félix (38° 58' 20" W, 12° 36' 17" S), São Felipe (39° 05' 22" W, 12° 50' 50" S) e Maragogipe (38° 55' 10" W, 12° 46' 40" S), localizados na região do Recôncavo Baiano. O solo predominante dessas localidades é o latossolo, o clima é úmido e subúmido, com temperatura anual de 23,4°C, exceto São Felipe, que apresenta 24,4°C, e a precipitação média anual varia de 1087,7 a 1173,9 mm (SEI, 2013).

A identificação dos genótipos foi formulada por códigos com três letras, sendo que a primeira representa a cultura (I – inhame) e as subsequentes, o município no qual o genótipo foi coletado (CA - Cruz das Almas; FL – São Felipe; FX – São Félix e MA – Maragogipe, seguido pelo número do genótipo (Tabela 1).

Tabela 1 - Relação dos genótipos de inhame (*Dioscorea rotundata*) coletados na Região do Recôncavo Baiano com seus respectivos códigos, série, procedência (povoado e município) e número. Cruz das Almas, BA. 2021.

Códigos dos Genótipos	Série dos Genótipos	Procedência		Nº de Genótipos
		Povoado	Município	
ICA	1 a 12	Tua	Cruz das Almas	12
ICA	13 a 18	Cadete	Cruz das Almas	6
ICA	19 a 24	Combé	Cruz das Almas	6
ICA	25 a 30	Três Boca	Cruz das Almas	6
IFL	1 a 6	Camargo	São Felipe	6
IFL	7 a 12	Jaracandá	São Felipe	6
IFL	13 a 18	Bom Gosto	São Felipe	6
IFL	19 a 24	Boa Esperança	São Felipe	6
IFL	25 a 30	Campo das Flores	São Felipe	6
IFX	1 a 6	São Bento	São Félix	6
IFX	7 a 12	Engenho de São João	São Félix	6
IFX	13 a 18	Monte Alegre	São Félix	6
IFX	19 a 24	Boa Vista	São Félix	6
IFX	25 a 30	Matatauba	São Félix	6
IMA	1 a 6	Serraria	Maragogipe	6
IMA	7 a 12	Campinas	Maragogipe	6
IMA	13 a 18	Encruzilhada	Maragogipe	6
IMA	19 A 30	Batatans	Maragogipe	12
Total	-	-	-	120

I- Inhame; CA - Cruz das Almas; FL – São Felipe; FX – São Félix e MA – Maragogipe.

Realizou-se a caracterização dos genótipos em duas etapas. A primeira etapa ocorreu entre os meses de maio e junho de 2015, em que foram realizadas quatro visitas nos povoados dos municípios citados na Tabela 1. Em cada propriedade rural, foram selecionados aleatoriamente seis genótipos, os quais foram marcados com uma fita com a respectiva identificação de 1 a 6, sendo caracterizada a parte aérea (folhas e caules), com o auxílio de régua e paquímetro.

A segunda etapa ocorreu no mês de dezembro de 2015, em que foram realizadas duas visitas nas quais foram coletados tubérculos dos genótipos previamente marcados, ensacados e transportados para o Laboratório de Tecnologia de Alimentos da Universidade Federal do Recôncavo da Bahia (UFRB), onde foram avaliados com auxílio de fita métrica e balança.

2.2 Redes Neurais Artificiais

Para o treinamento das redes, as variáveis de entrada (ou independentes) foram o diâmetro do caule (DC), comprimento do pecíolo (CP), distância entre a inserção do pecíolo na folha a extremidade superior da folha (P3), distância entre a inserção do pecíolo na folha a extremidade inferior da folha (P2) e a largura da folha (F1), e a variável de saída (ou dependente) foi o peso do tubérculo.

Segundo Haykin (2001), a camada de entrada é responsável por receber os valores (quantitativos ou qualitativos) das variáveis apresentadas e transmitir para a camada intermediária ou oculta, enquanto que as camadas intermediária e de saída mapeiam o conhecimento e processam as informações com os seus neurônios, os quais também são chamados de nós. Dessa forma, os neurônios recebem os sinais de entrada multiplica com seus respectivos pesos e adiciona um sinal preconizado (*bias*). A resposta atua mediante a uma função de ativação e, assim, é fornecida a saída de neurônio (Figura 1).

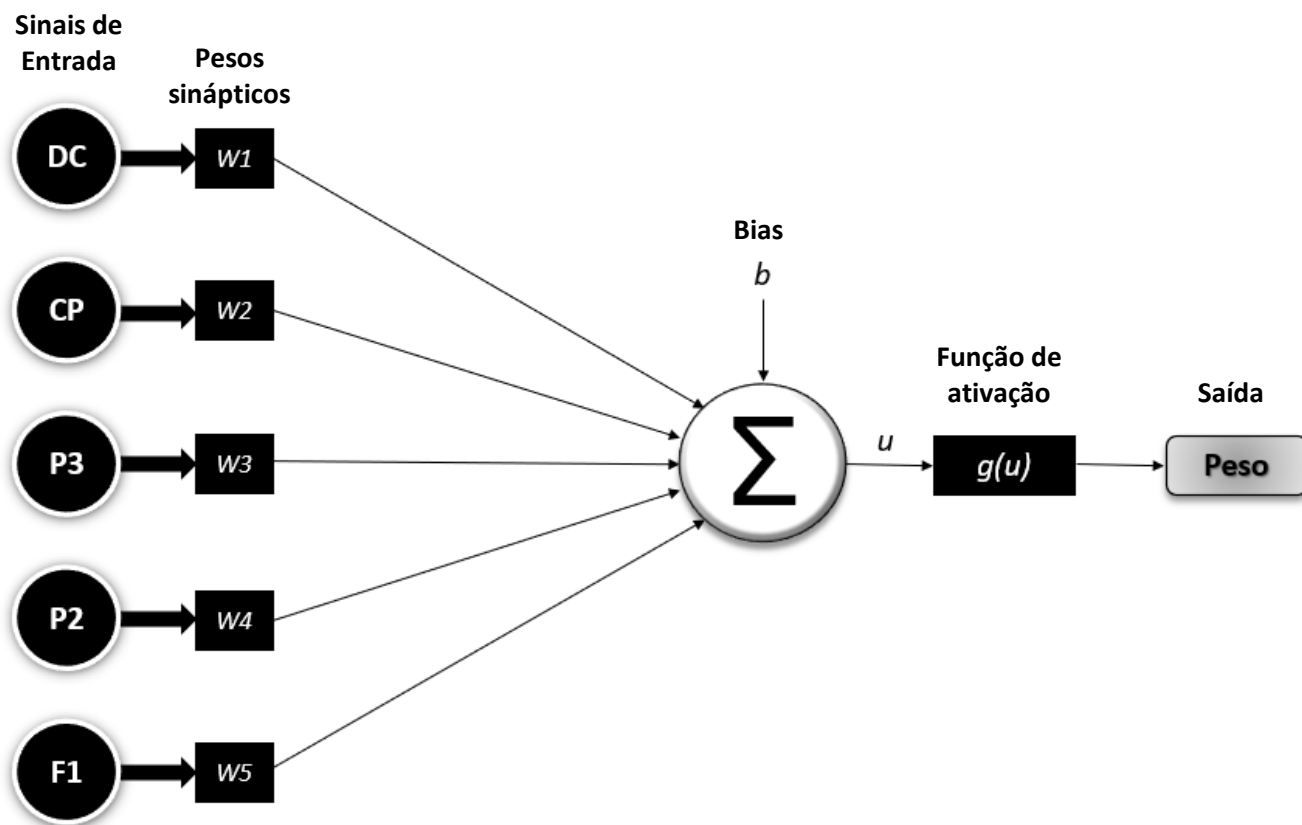


Figura 1 - Representação do neurônio artificial dentro de uma rede do tipo Perceptron de Múltiplas Camadas ou *Multilayer Perceptron* (MLP). DC: diâmetro do caule; CP: comprimento do pecíolo; P3: distância entre a inserção do pecíolo na folha a extremidade superior da folha; P2: distância entre a inserção do pecíolo na folha a extremidade inferior da folha; F1: largura da folha.

Fonte: Arquivo Pessoal (2021).

O treinamento de uma RNA começa quando ocorre a apresentação dos dados (variáveis de entrada e saída) a uma arquitetura pré-estabelecida. Esse processo inicia com valores aleatórios dos pesos sinápticos, e a partir disso, os valores estimados da primeira saída é comparada com os valores observados da variável resposta. A diferença entre o valor estimado pela rede e o valor observado gera um sinal de erro que calibra o ajuste de pesos, e assim, inicia um novo ciclo, a fim de aproximar o valor estimado da variável resposta do resultado desejado, com intuito de minimizar o erro (HAYKIN, 2001).

Neste estudo, utilizou-se a arquitetura do tipo Perceptron de Múltiplas Camadas ou *Multilayer Perceptron* (MLP) nas redes treinadas, amplamente usada para desenvolver o modelo de predição, a qual consiste em uma camada de entrada com neurônios representando variáveis independentes, uma camada de saída com neurônio representando a variável dependente e uma ou mais

camadas ocultas contendo neurônio para ajudar a capturar a não linearidade do sistema.

Para garantir uma modelagem MLP com êxito, dois princípios importantes devem ser considerados, são eles: o número de camadas ocultas e o número de neurônios em cada camada oculta. Como a maioria dos problemas na modelagem de redes neurais artificiais podem ser resolvidos com uma camada oculta, utilizou-se, nesta pesquisa, uma camada oculta com neurônios variando de 1 a 5, adotando-se o número de épocas (ou ciclos) de 100 a 1000. Além disso, testou-se as funções de ativação linear, sigmoideal, ReLu, identidade e tangente hiperbólica.

Os dados foram divididos aleatoriamente em duas partes, sabendo do papel importante da divisão das amostras para avaliação do desempenho de uma RNA, dividiu-se 75% para o treinamento e 25% para a validação. Após a separação dos dados, foram calculadas as estatísticas descritivas das variáveis quantitativas: valores mínimo e máximo, média, mediana, desvio padrão e coeficiente de variação. E, para obter as RNA, utilizou-se o pacote *h2o* (LEDELL et al., 2020) do *software* RStudio versão 1.4.1103 (RSTUDIO TEAM, 2020).

Sendo assim, treinou-se um total de 1000 RNA, e destas, selecionou-se as cinco melhores para serem apresentadas. Desse modo, avaliou-se a qualidade das RNA nas etapas de treino e validação mediante as seguintes estatísticas:

- Raiz do Erro Quadrático Médio (REQM) da variável estimada: Avalia a diferença quadrática média entre os valores observados e os valores estimados. Quanto menor o REQM, melhor será a precisão da estimativa.

$$REQM = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

Sendo:

y_i : Valor observado da i -ésima variável

\hat{y}_i : Valor estimado da i -ésima variável

n : Tamanho da amostra

- Erro Absoluto Médio (EAM) da variável estimada: Permite medir a magnitude média do erro entre os valores observados e os valores estimados. A predição adequada e eficaz da variável estimada é definida pelo menor valor do EAM, ou seja, esse parâmetro de avaliação se aproxima do zero à medida que as discrepâncias entre a previsão e a observação diminuem, e é calculada da seguinte forma:

$$EAM = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$$

Sendo:

y_i : Valor observado da i -ésima variável

\hat{y}_i : Valor estimado da i -ésima variável

n : Tamanho da amostra

Além disso, fez-se gráficos de distribuição de resíduos, que consiste na verificação estatística da dispersão dos erros percentuais em relação aos valores observados e estimados, uma vez que mesmo que as estatísticas de qualidade de ajuste apresentadas sejam boas indicadores para a seleção dos modelos, a análise gráfica de resíduo é quesito fundamental na escolha de um modelo, considerando-se que erros de tendência podem ocorrer em determinada amplitude de classe da variável resposta e não serem detectados pelas demais estatísticas.

3. RESULTADOS

Os valores mínimos, médios, máximos e medianos das variáveis de entrada (ou independentes) e o de saída (dependente), assim como o desvio padrão e o coeficiente de variação (CV), podem ser observados na tabela 2. De acordo com Gomes (2000), a variabilidade de qualquer atributo pode ser classificada segundo a magnitude do seu coeficiente de variação. Se o CV for

inferior a 10%, considera-se de baixa variabilidade, de 10% a 20%, média, de 20% a 30%, alta e maior que 30%, muito alta. Portanto, as variáveis com maiores coeficientes de variação foram o peso do tubérculo e a distância entre a inserção do pecíolo na folha a extremidade superior da folha (P3) com 43,27% e 29,84%, nos dados de treinamento, e, 38,40% e 36,48%, nos dados de validação, respectivamente. Assim, foi possível observar que a variável de produção (o peso do tubérculo) apresentou uma variabilidade muito alta, expressa pelo seu coeficiente de variação. Esta variabilidade é fundamental para subsidiar os programas de melhoramento, uma vez que esta característica tem influência direta no valor comercial do inhame (AFONSO et al., 2014).

Tabela 2. Estatística das variáveis mensuradas nos genótipos de inhame (*Dioscorea rotundata*) caracterizados na Região do Recôncavo Baiano. Cruz das Almas, BA. 2021.

Variável	CONJUNTO DE DADOS DE TREINAMENTO						CONJUNTO DE DADOS DE VALIDAÇÃO					
	Mínimo	Média	Mediana	Máximo	Desvio Padrão	CV (%)	Mínimo	Média	Mediana	Máximo	Desvio Padrão	CV (%)
DC (cm)	0,2	0,426	0,4	0,6	0,099	23,25	0,3	0,427	0,4	0,6	0,087	20,35
CP (cm)	3	4,856	5	7,2	0,940	19,35	3	4,760	4,65	7,6	1,062	22,31
P3 (cm)	2,1	3,916	4	8	1,168	29,84	1,1	3,363	3,2	7,7	1,227	36,48
P2 (cm)	9,3	14,354	14,25	19,8	2,137	14,89	8,5	13,757	13,85	17,5	2,059	14,97
F1 (cm)	4	6,570	6,2	11,3	1,618	24,63	4	6,347	6,2	9,9	1,721	27,12
Peso (kg)	0,318	0,800	0,721	2,005	0,346	43,27	0,37	0,804	0,725	1,599	0,309	38,40

CV – Coeficiente de Variação; DC – Diâmetro do caule; CP – Comprimento do pecíolo; P3 – Distância entre a inserção do pecíolo na folha a extremidade superior da folha; P2 – Distância entre a inserção do pecíolo na folha a extremidade inferior da folha e F1 – Largura da folha.

As redes selecionadas exibem a seguinte estrutura: cinco neurônios na camada de entrada, um neurônio na camada oculta e um neurônio na camada de saída que representa o peso do tubérculo (Tabela 3). O número de neurônios na camada oculta está diretamente relacionado com a capacidade da RNA detectar relações não lineares entre as informações apresentadas em cada camada de entrada e saída e, assim, obter estatísticas satisfatórias. Logo, o número excessivo de neurônios nesta camada pode ocasionar a memorização dos dados pela rede durante o processo de treinamento, resultando na perda de sua capacidade de generalização. Por outro lado, um pequeno número de neurônios na camada oculta pode não ser suficiente para a realização da tarefa

desejada. Dessa forma, o mais indicado é escolher configurações mais simples com intuito de obter resultados adequados (LEAL; MIGUEL; MATRICARDI, 2020).

Tabela 3. Redes Neurais Artificiais treinadas para estimativa do peso do tubérculo do inhame (*Dioscorea rotundata*) em função de caracteres agronômicos na Região do Recôncavo Baiano. Cruz das Almas, BA. 2021.

Rede Neural Artificial	Arquitetura	REQM		EAM	
		TREINAMENTO	VALIDAÇÃO	TREINAMENTO	VALIDAÇÃO
RNA 1	5-1-1	0,265	0,248	0,220	0,218
RNA 2	5-1-1	0,231	0,299	0,178	0,232
RNA 3	5-1-1	0,217	0,270	0,158	0,220
RNA 4	5-1-1	0,263	0,316	0,214	0,259
RNA 5	5-1-1	0,269	0,307	0,219	0,250

RQEM – Raiz do Erro Quadrático Médio; EAM – Erro Absoluto Médio.

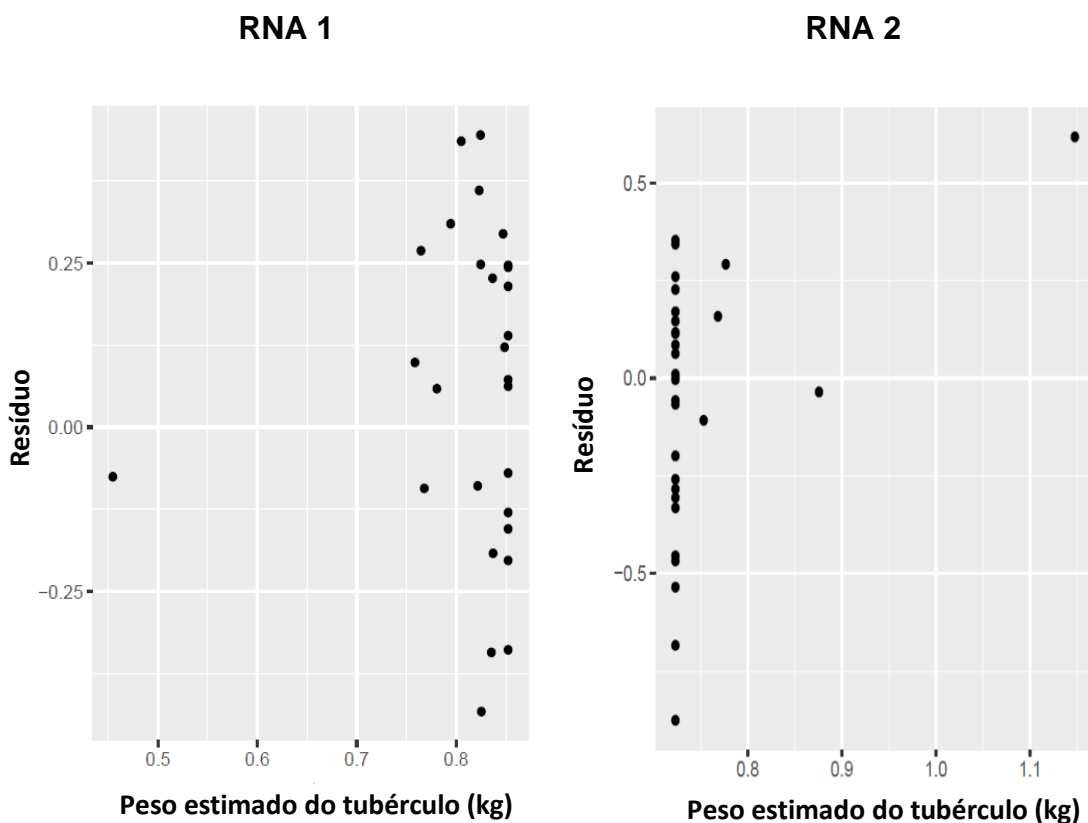
Para avaliar o desempenho das RNA, os parâmetros de avaliação (Raiz do Erro Quadrático Médio (REQM) e Erro Absoluto Médio (EAM)) foram utilizados para expressar a precisão do modelo juntamente com a análise residual. Uma vez que se faz necessário a combinação desses indicadores para determinar e/ou comparar a eficiência geral dos modelos. Dentre as RNA selecionadas (Tabela 3), a que apresentou os menores valores de REQM e EAM na validação foi a RNA 1 (REQM = 0,248; EAM = 0,218), enquanto os maiores se apresentaram na RNA 4 (REQM = 0,316; EAM = 0,259). As diferenças entre a REQM e o EAM nas etapas de treinamento e de validação foram baixas, indicando que os processos de calibração e generalização forneceram uma resposta satisfatória da rede. Portanto, para RNA, a validação é um procedimento fundamental, pois, de acordo com os resultados obtidos, nem sempre a melhor RNA do treinamento é a mais precisa e eficaz (VENDRUSCOLO et al., 2015). Por isso, o procedimento de validação deve ser aplicado para verificar a capacidade de uma rede neural produzir saídas adequadas para os valores de entradas que não estavam presentes durante a etapa de treinamento (BINOTI et al., 2015). Assim, essas duas capacidades de processamento de informação tornam possível a resolução de problemas complexos (ALMEIDA, 2002).

Embora, a REQM e o EAM sejam critérios estatísticos de bons indicadores para a seleção de uma rede, visto que permitem medir o desempenho global do

ajuste, e quanto menores as estimativas, mais confiável é o modelo ajustado, os mesmos devem ser associados com outros índices para uma melhor avaliação. Por isso, foi utilizado a análise gráfica residual para a confirmação desses dados.

A análise gráfica de resíduos em função do peso estimado do tubérculo do inhame das redes que melhor se ajustaram à predição está presente na Figura 2. Mediante a esta análise, foi possível observar que as melhores redes possuíram uma leve tendenciosidade, exceto a RNA1, a qual apresentou seus resíduos distribuídos de forma homogênea.

A RNA1 se destacou de maneira satisfatória apresentando um bom ajuste e ausência de *outliers*, isto é, pontos distantes da reta guia em que o ajuste não se deu de maneira eficiente. Assim, ao comparar a análise gráfica residual de todas as cinco RNA, nota-se que a predição do peso do tubérculo do inhame foi mais eficiente na RNA 1, corroborando com os resultados dos parâmetros de avaliação de desempenho das RNA (REQM e EAM).



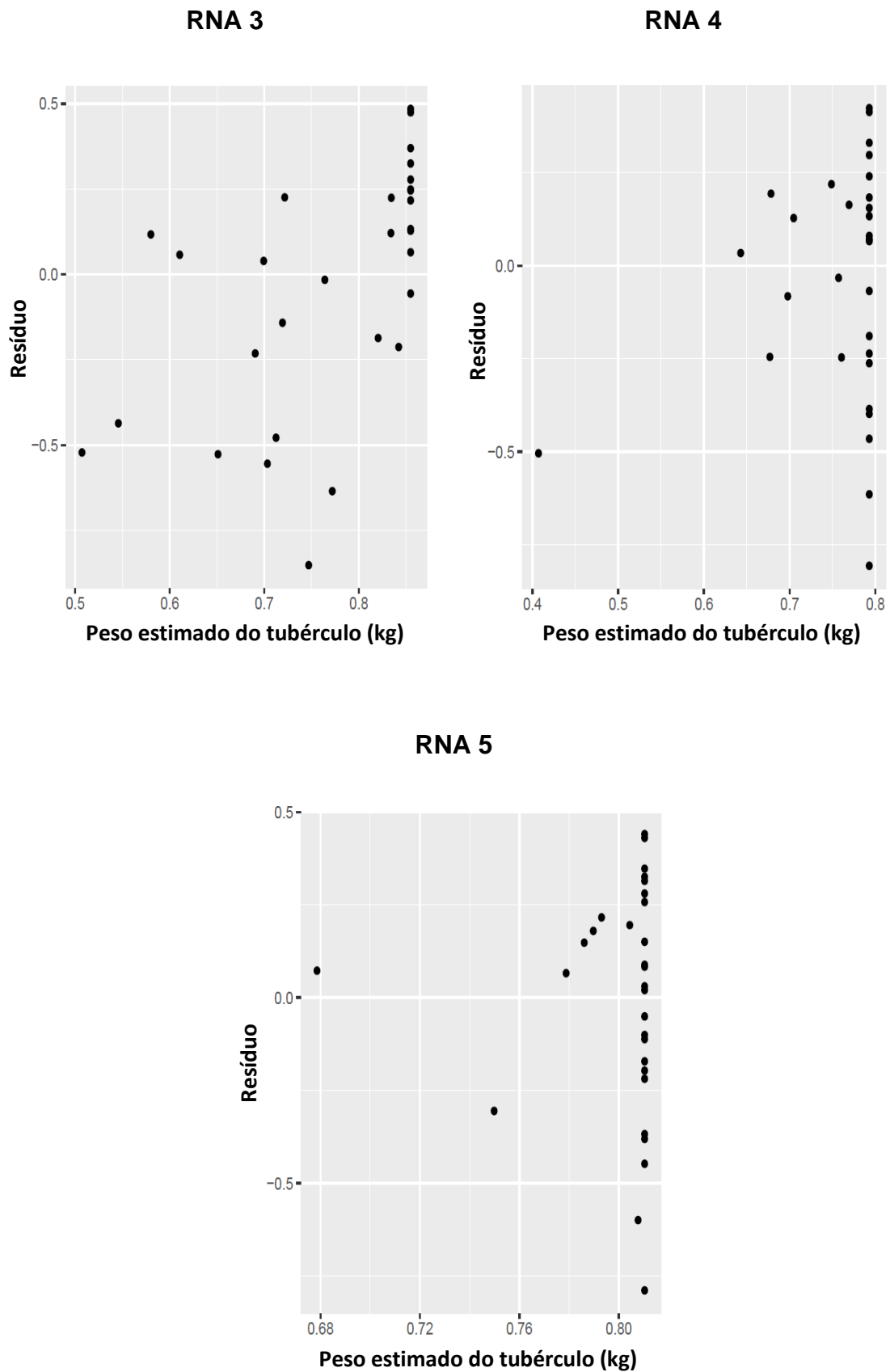


Figura 2. Gráfico de resíduos em função do peso estimado do tubérculo do inhame das cinco melhores redes neurais artificiais treinadas na etapa de validação.

Fonte: Arquivo Pessoal (2021).

4. DISCUSSÃO

4.1 IMPORTÂNCIA DA VIABILIZAÇÃO DA CULTURA DE INHAME NA BUSCA DA PRODUTIVIDADE

O sucesso das safras de inhame está diretamente relacionado ao planejamento do processo produtivo da cultura, à qualidade dos tubérculos e insumos, aos tratos culturais, às questões climáticas, aos preços de comercialização e aos custos de produção, para uma correta tomada de decisões (TOLOI et al., 2021). Desta forma, o setor agrícola busca atingir uma estimativa da produção antes do período da colheita, visto que poderá utilizar a previsão da safra para analisar as eventuais necessidades de transporte e armazenamento do produto, como também os possíveis lucros na sua comercialização (KHAIRUNNIZA-BEJO; MUSTAFFHA; ISMAIL, 2014). Com isso, o planejamento das atividades agrícolas visa minimizar perdas relativas ao padrão comercial dos tubérculos, as quais estão associadas à qualidade mercadológica (exigência do mercado por um determinado padrão), decorrentes a distúrbios fisiológicos, ao metabolismo fisiológico pós-colheita (perdas de peso e brotações), e a danos mecânicos e aos ocasionados por pragas e doenças (KIAYA, 2014; OSUNDE, 2008).

Os estabelecimentos comerciais valorizam a aparência dos produtos, visando os compatíveis aos atributos de padrão de qualidade. Ou seja, os rizomas considerados pequenos, com defeitos e que não foram devidamente higienizados (eliminação das raízes e da terra aderida à sua superfície) apresentam baixo valor comercial ou não têm comercialização em determinados mercados (GARRIDO et al., 2017). Enfim, os melhores preços são obtidos dos tubérculos com mais de 100g, limpos e sem deformidades (CEASAMINAS, 2021).

O inhame desidrata facilmente, o que ocasiona significativa perda no quantitativo do produto a ser comercializado. Logo, o planejamento, a cura e as condições adequadas de armazenamento e transporte são procedimentos indispensáveis para a minimização da perda de peso e para o desenvolvimento do padrão de qualidade dos tubérculos do inhame (KIAYA, 2014).

De acordo com Amusa et al. (2003), o armazenamento em condições de altas temperaturas e pouca ventilação predispõe à maior perda de peso. Maalekuu, Saajah e Addae (2014) investigaram as melhores condições de armazenamento que prolongam e mantêm as qualidades comestíveis e comercializáveis, e a vida útil de alguns cultivares de inhame branco (*Dioscorea rotundata*). Os resultados mostraram que houve uma diferença significativa entre as cultivares Pona e Tela, onde ambas tiveram 21,12% e 25,95% de perda de peso num período de 150 dias, respectivamente. Assim, a cultivar Tela é mais propensa a perder peso. Entretanto, os métodos de armazenamento (celeiro tradicional, estrutura de armazenamento aberto e armazenamento de pilha no chão) não apresentaram efeito significativo na redução da perda de peso, pois apresentaram desempenho estatisticamente igual para minimizar a perda de peso entre os tubérculos armazenados.

A perda de peso é uma das manifestações mais graves de degradação do tubérculo do inhame, que pode ser devido a reações deletérias (MAALEKUU; SAAJAH; ADDAE, 2014). Dentre os diversos aspectos relacionados à perda de peso dos tubérculos, uma das causas é o déficit de pressão de vapor, que é diferença entre a pressão de vapor do produto e a pressão de vapor do ar circundante. Isto significa que quando a pressão de vapor d'água no tubérculo é maior que a do ambiente, ocorre a perda de água para o meio (TAIZ; ZEIGER, 2004). Associado a essa circunstância de fundamento fisiológico, existem ainda os danos mecânicos, a ação de patógenos ou a combinação desses, que podem ocasionar as perdas de peso da matéria fresca durante a fase pós-colheita (OSUNDE, 2008; FAO, 2021). Enfim, a magnitude da perda de água é um princípio crítico para a conservação dos tubérculos, podendo representar a maior porcentagem da perda de sua matéria fresca, uma vez que a respiração excessiva (em grande parte devido à oxidação do amido armazenado), outro fator da perda de peso total, geralmente, proporciona apenas de 3 a 5% da perda de peso (FAO, 2021).

4.2 DESENVOLVIMENTO E UTILIZAÇÃO DA MODELAGEM NA AGRICULTURA

No cenário agrícola atual, o desenvolvimento e a utilização de modelagem são relevantes na previsão de safra e/ou exploração do potencial produtivo da cultura, uma vez que, mediante as informações geradas por sistemas inteligentes artificiais, o produtor poderá monitorar sua produção de forma mais efetiva, obtendo um alto índice de desempenho de mercado e competitividade (SOUSA et al., 2020). Assim, modelos de predição podem contribuir para a análise de dados e tomada de decisões no manejo de uma cultura, permitindo uma avaliação precoce da produtividade das culturas para melhorar o planejamento e gestão agrícola, bem como a otimização do tempo e a aplicação adequada dos recursos (SOARES et al., 2015).

O desenvolvimento de modelos de predição baseados em caracteres agronômicos utilizando redes neurais artificiais (RNA) foi capaz de prever a resposta da produção de frutos de 692 genótipos de pimenta (*Capsicum annuum* L.) a partir da análise de oito caracteres da planta (GHOLIPOOR; NADALI, 2019). Soares et al. (2013) estabeleceram uma metodologia baseada em estudos com RNA que permitiu a predição do peso do cacho em bananeiras cv. Tropical, o qual foi estimado com precisão e eficiência. Rad et al. (2015) também propuseram uma metodologia com aplicação de RNA para prever o peso do fruto do melão (*Cucumis melo* L.), além de identificar com eficácia e acurácia o melhor preditor. Alves et al. (2018) demonstraram que é possível estimar a produtividade da soja (*Glycine max* (L.) Merrill) baseando-se nas principais características agronômicas com diferentes hábitos de crescimento e submetidas a diferentes densidades de semeadura.

Deste modo, pesquisas fundamentadas em caracteres fenotípicos de interesse em programas de melhoramento, que podem apresentar aumento de produtividade na cultura, são consideradas relevantes para identificação e seleção de indivíduos superiores, visto que podem ser recomendados para incorporação nos sistemas de produção dos agricultores da região (HESLOP-HARRISON; SCHWARZACHER, 2007). Esses caracteres são geralmente quantitativos, permitindo visualizar de forma preliminar a adaptação e o potencial

produtivo dos fenótipos, constituídos por baixa herdabilidade, desejáveis sob ponto de vista econômico e a expressão gênica é influenciada pelas condições ambientais (ELAMEEN et al., 2011; LAURIE et al., 2013). Sendo assim, o conhecimento dessas características, a partir de modelos preditivos, pode facilitar decisões antecipadas por parte do agricultor e/ou melhorista, além de possibilitar ao produtor fazer uma estimativa das projeções para a safra e prever aspectos de financiamento antes da colheita (SOARES et al., 2015).

Apesar da cultura do inhame apresentar uma importância socioeconômica significativa, poucos estudos voltados a aplicabilidade das RNA como ferramentas preditivas foram realizados. Portanto, a falta de informação associada com a intensa concorrência pela obtenção de recursos para pesquisa justificam a utilização de modelagem para a predição, uma vez que há economia de tempo, trabalho e quantidade de recursos necessários para o planejamento e tomada de decisões de manejo no setor agrícola, além de que a formulação de um modelo convencional é difícil e limitado, devido à complexidade e a não linearidade nos dados dos parâmetros relacionados a ela, requerendo informações detalhadas sobre os processos físicos do sistema a ser modelado (SUDHEER; GOSAIN; RAMASASTRI, 2003).

Embora as RNA encontrem dificuldades no treinamento, como: definir a melhor arquitetura (número de camadas da rede treinada e a quantidade de neurônios presentes nas camadas ocultas), as funções de ativação que estão presentes tanto nas camadas ocultas quanto nas camadas de saída (MARTINS et al., 2016), e a determinação do número ideal de ciclos de treinamento da rede (BASHEER; HAJMEER, 2000), uma vez que pode propiciar a subestimação (*underfitting*) ou a superestimação (*overfitting*) dos dados (JABBAR; KHAN, 2015), o emprego das redes neurais tem se destacado promissora em virtude da possibilidade de um desempenho superior aos modelos convencionais utilizados na solução de problemas (BRAGA; CARVALHO; LUDERMIR, 2007).

4.3 APRIMORAMENTO DA MODELAGEM PARA PREDIÇÃO NA CULTURA DO INHAME

Apesar dos resultados promissores da utilização do sistema de Redes Neurais Artificiais para predição do peso do tubérculo do inhame, estudos futuros podem ser conduzidos na perspectiva da aplicação de dados coletados em outras áreas experimentais bem como na utilização de diversas variáveis de interesse nas lavouras, na otimização dos hiperparâmetros, por exemplo, a taxa de aprendizado, o parâmetro de regularização, número de camadas e neurônios em cada uma delas, a variação da quantidade de dados nos parâmetros de entrada, as funções de ativação, dentre outros, uma vez que o desempenho e a capacidade de aprendizado da técnica estão diretamente associados a escolha adequada da arquitetura da rede. Quanto aos atributos do neurônio, como: pesos e *bias*, é possível averiguar, mediante a cálculos, suas reais influências no problema proposto variando seus valores e buscando a redução do erro quadrático médio.

Por fim, destaca-se que a proposta do emprego da modelagem em predição não é substituir a experimentação convencional, nem se deseja obter uma única e exclusiva resposta para questões de natureza complexa e mutável. O que se pretende é tornar a atividade científica mais eficiente e, conseqüentemente, alcançar os resultados desses trabalhos de modo menos oneroso e dispendioso economicamente, além de menos laborioso, visto que servem para a orientação de pesquisas, gestão de tecnologia e tomada de decisão.

5. CONCLUSÃO

- 1) As Redes Neurais Artificiais (RNA) se apresentaram como uma ferramenta satisfatória na predição do peso do tubérculo do inhame, demonstrando que podem fornecer praticidade e agilidade no processamento de dados da cultura, sem que perca a confiabilidade nas informações apresentadas.

- 2) A arquitetura MLP 5-1-1 foi a mais indicada para a estimativa do peso do tubérculo do inhame na Região do Recôncavo Baiano.
- 3) A validação realizada confirmou a eficácia da RNA aplicada aos dados, o que tornou o método bastante eficiente na predição do peso do tubérculo do inhame, uma vez que foi verificado que nem sempre a melhor RNA obtida no treino apresentará um bom desempenho quando aplicada em outro conjunto de dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- AFONSO, S. D. J.; MOREIRA, R. F. C.; CONCEIÇÃO, L. S.; SILVA, S. A. Agronomic characterization of yam genotypes production. **IOSR Journal of Agriculture and Veterinary Science**, v. 7, n. 12, p. 22-30, 2014.
- ALMEIDA, J. S. Predictive non-linear modeling of complex data by artificial neural networks. **Current opinion in biotechnology**, v. 13, n. 1, p. 72-76, 2002.
- ALVES, G. R.; TEIXEIRA, I. R.; MELO, F. R.; SOUZA, R. T. G.; SILVA, A. G. Estimating soybean yields with artificial neural networks. **Acta Scientiarum. Agronomy**, v. 40, 2018.
- AMUSA, N. A.; ADIGBITE, A. A.; MUHAMMED, S.; BAIYEWU, R. A. Yam diseases and its management in Nigeria. **African Journal of Biotechnology**, v. 2, n. 12, p. 497-502, 2003.
- ASIEDU, R.; SARTIE, A. Crops that feed the world 1. Yams. **Food Security**, v. 2, n. 4, p. 305-315, 2010.
- BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B.; GLERIANI, J. M. Prognose em nível de povoamento de clones de eucalipto empregando redes neurais artificiais. **Cerne**, v. 21, p. 97-105, 2015.
- CENTRAIS DE ABASTECIMENTO DE MINAS GERAIS S.A. (CEASAMINAS). Informações Nutricionais Inhame. Disponível em: <<http://ceasa.mg.gov.br/informacoesnutricionaisprodutogeral.asp?ComboProdutos=INH>>. Acesso em: 08 Set. 2021.
- ELAMEEN, A.; LARSEN, A.; KLEMSDAL, S. S.; FJELLHEIM, S.; SUNDHEIM, L.; MSOLLA, S.; MASUMBA, E.; ROGNLI, O. A. Phenotypic diversity of plant morphological and root descriptor traits within a sweet potato, *Ipomoea batatas* (L.) Lam., germplasm collection from Tanzania. **Genetic Resources and Crop Evolution**, v. 58, n. 3, p. 397-407, 2011.
- EPPING, J.; LAIBACH, N. An underutilized orphan tuber crop—Chinese yam: a review. **Planta**, v. 252, n. 4, p. 1-19, 2020.
- FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO). Factors affecting storability of roots and tubers. Disponível em: <<https://www.fao.org/3/x5415E/x5415e02.htm>>. Acesso em: 08 Set. 2021.
- FARJAM, A.; OMID, M.; AKRAM, A.; NIARI, Z. F. A neural network based modeling and sensitivity analysis of energy inputs for predicting seed and grain corn yields. **Journal of Agricultural Science and Technology**, v. 16, n. 4, p. 767-778, 2014.
- FORTIN, J. G.; ANCTIL, F.; PARENT, L.; BOLINDER, M. A. A neural network experiment on the site-specific simulation of potato tuber growth in Eastern

Canada. **Computers and Electronics in Agriculture**, v. 73, n. 2, p. 126-132, 2010.

GARRIDO, M. S.; SOARES, A. C. F.; MENEZES, R. S. C.; SAMPAIO, E. V. S. B.; SOUZA, E. M. S.; MACEDO, B. L. N. Quality and yield of yam tubers cultivated with green manure in the Northeastern Brazil. **Comunicata Scientiae**, v. 8, n. 2, p. 275-280, 2017.

GHOLIPOOR, M.; NADALI, F. Fruit yield prediction of pepper using artificial neural network. **Scientia Horticulturae**, v. 250, p. 249-253, 2019.

GOMES, F. P. **Curso de estatística experimental**. 14 ed. Piracicaba: Degaspari, 2000. 477 p.

GUIMARÃES, B. V. C.; DONATO, S. L. R.; ASPIAZÚ, I.; AZEVEDO, A. M. Yield prediction of 'Prata Anã' and 'BRS Platina' banana plants by artificial neural networks. **Pesquisa Agropecuária Tropical**, v. 51, 2021.

HAYKIN, S. **Redes Neurais: Princípios e Práticas**. Porto Alegre: Bookman, 2001. 900 p.

HESLOP-HARRISON, J. S.; SCHWARZACHER, T. Domestication, genomics and the future for banana. **Annals of botany**, v. 100, n. 5, p. 1073-1084, 2007.

INTERNATION BUSINESS MACHINES CORPORATION (IBM). Neural Networks. Disponível em: <<https://www.ibm.com/cloud/learn/neural-networks>>. Acesso em: 27 Jun. 2021.

KIAYA, V. Post-harvest losses and strategies to reduce them. **Technical Paper on Postharvest Losses, Action Contre la Faim (ACF)**, v. 25, 2014.

LAURIE, S. M.; CALITZ, F. J.; ADEBOLA, P. O.; LEZAR, A. Characterization and evaluation of South African sweet potato (*Ipomoea batatas* (L.) LAM) land races. **South African Journal of Botany**, v. 85, p. 10-16, 2013.

LEAL, F. A.; MIGUEL, E. P.; MATRICARDI, E. A. T. Estimates of deforestation rates in rural properties in the legal amazon. **Floresta e Ambiente**, v. 27, 2020.

LEDELL, E.; GILL, N.; AIELLO, S.; FU, A.; CANDEL, A.; CLICK, C.; KRALJEVIC, T.; NYKODYM, T.; ABOYOUN, P.; KURKA, M.; MALOHLAVA, M. R interface for the 'H2O' scalable machine learning platform. 2020. Disponível em: <<https://cran.r-project.org/web/packages/h2o/h2o.pdf>>. Acesso em: 31 Mai. 2021.

MAALEKUU, B. K.; SAAJAH, J. K.; ADDAE, A. K. Effect of three storage methods on the quality and shelf-life of white yam (*Dioscorea rotundata*) cultivars pona and tela. **Journal of Agricultural science**, v. 6, n. 7, p. 221, 2014.

MOVAGHARNEJAD, K.; NIKZAD, M. Modeling of tomato drying using artificial neural network. **Computers and electronics in agriculture**, v. 59, n. 1-2, p. 78-85, 2007.

MURASE, H. Artificial intelligence in agriculture. **Computers and Electronics in Agriculture**, v. 29, n. 1/2, 2000.

OSUNDE, Z. D. Minimizing Postharvest Losses in Yam (*Dioscorea* spp.): Treatments and Techniques. In: ROBERTSON, G. L.; LUPIEN, J. R. (org.). **Using Food Science and Tecnologu to Improve Nutrition and Promote National Development**. Oakville, Ontário, 2008. 169 p.

PFEIFER, R.; DAMIAN, D.; FUCHSLIN, R. Neural Networks. University of Zurich, Zurich, Suíça, 2010. Disponível em: <<https://www.ifi.uzh.ch/en.html>>. Acesso em: 27 Set. 2021.

RAD, M. R. N.; KOOHKAN, S.; FANAEI, H. R.; RAD, M. R. P. Application of Artificial Neural Networks to predict the final fruit weight and random forest to select important variables in native population of melon (*Cucumis melo* L.). **Scientia Horticulturae**, v. 181, p. 108-112, 2015.

RSTUDIO TEAM. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, 2020. Disponível em: <<https://www.rstudio.com/>>.

SOARES, F. C.; ROBAINA, A. D.; PEITER, M. X.; RUSSI, J. L. Predição da produtividade da cultura do milho utilizando rede neural artificial. **Ciência Rural**, v. 45, p. 1987-1993, 2015.

SOARES, J. D. R.; PASQUAL, M.; LACERDA, W. S.; SILVA, S. O.; DONATO, S. L. R. Utilization of artificial neural networks in the prediction of the bunches' weight in banana plants. **Scientia Horticulturae**, v. 155, p. 24-29, 2013.

SOUSA, A. M. C. B.; SILVA, V. B.; LOPES, A. C. A.; GOMES, R. L. F.; CARVALHO, L. C. B. Prediction of grain yield, adaptability, and stability in landrace varieties of lima bean (*Phaseolus lunatus* L.). **Crop Breeding and Applied Biotechnology**, v. 20, n. 1, 2020.

SUDHEER, K. P.; GOSAIN, A. K.; RAMASASTRI, K. S. Estimating actual evapotranspiration from limited climatic data using neural computing technique. **Journal of irrigation and drainage engineering**, v. 129, n. 3, p. 214-218, 2003.

SUPERINTENDÊNCIA DE ESTUDOS ECONÔMICOS E SOCIAIS DA BAHIA (SEI). **Estatísticas dos Municípios Baianos**. Salvador: SEI, v. 4, n. 1, 2013. p. 1-388. Disponível em: <https://www.sei.ba.gov.br/index.php?option=com_content&view=article&id=2441&Itemid=284>. Acesso em: 31 Mai. 2021.

TAIZ, L.; ZEIGER, E. **Fisiologia Vegetal**. 4. ed. Porto Alegre: Artmed, 2004. 849p.

TOLOI, R. C.; REIS, J. G. M.; TOLOI, M. N. V.; VENDRAMETTO, O.; CABRAL, J. A. S. P. Applying analytic hierarchy process (AHP) to identify decision-making in soybean supply chains: a case of Mato Grosso production. **Revista de Economia e Sociologia Rural**, v. 60, 2021.

VENDRUSCOLO, D. G. S.; DRESCHER, R.; SOUZA, H. S.; MOURA, J. P. V. M.; MAMORÉ, F. M. D.; SIQUEIRA, T. A. S. Estimativa da altura de eucalipto por meio de regressão não linear e redes neurais artificiais. **Revista Brasileira de Biometria**, v. 33, n. 4, p. 556-569, 2015.

CAPÍTULO II

REDES NEURAIS ARTIFICIAIS NA ANÁLISE DE DIVERGÊNCIA GENÉTICA EM INHAME (*Dioscorea rotundata*)

REDES NEURAIS ARTIFICIAIS NA ANÁLISE DE DIVERGÊNCIA GENÉTICA EM INHAME (*Dioscorea rotundata*)

Autor: Ingrid Giovanna Vieira Santos

Orientador: Prof. Dr. Ricardo Franco Cunha Moreira

Coorientador: Prof. Dr. Liniker Fernandes da Silva

RESUMO: O objetivo deste estudo foi estimar e organizar a divergência genética dos genótipos de inhame (*Dioscorea rotundata*) mediante a análise dos Mapas Auto-Organizáveis de Kohonen. Para isso, foram utilizados os dados da caracterização agrônômica de 120 genótipos, os quais foram caracterizados com base em seis descritores quantitativos estabelecidos pelo IPGRI/IITA (1997): diâmetro do caule, em cm (DC), comprimento do pecíolo, em cm (CP), distância entre a inserção do pecíolo na folha a extremidade superior da folha, em cm (P3), distância entre a inserção do pecíolo na folha a extremidade inferior da folha, em cm (P2), largura da folha, em cm (F1) e peso do tubérculo, em kg (PT). Para reconhecer a organização da dissimilaridade dos genótipos de inhame, os Mapas Auto-Organizáveis de Kohonen (SOM) foram utilizados. E, para realizar a análise, utilizou-se o pacote *Kohonen* do software RStudio versão 1.4.1103. Assim, os resultados do estudo indicaram que a aplicação da inteligência artificial mediante a análise do Mapa Auto-Organizável de Kohonen foi eficaz para estudos de divergência genética em inhame da Região do Recôncavo Baiano.

PALAVRAS-CHAVE: Inteligência artificial; Mapa auto-organizável de Kohonen; Melhoramento genético.

ARTIFICIAL NEURAL NETWORKS IN THE ANALYSIS OF GENETIC DIVERGENCE IN YAM (*Dioscorea rotundata*)

Author: Ingrid Giovanna Vieira Santos

Advisor: Prof. Dr. Ricardo Franco Cunha Moreira

Co-Advisor: Prof. Dr. Liniker Fernandes da Silva

ABSTRACT: The aim of this study was to estimate and organize the genetic divergence of yam (*Dioscorea rotundata*) genotypes through the analysis of Kohonen's Self-Organizing Maps. For this, data from the agronomic characterization of 120 genotypes were used, which were characterized based on six quantitative descriptors established by IPGRI/IITA (1997): stem diameter, in cm (DC), petiole length, in cm (CP), distance between the insertion of the petiole in the leaf and the upper edge of the leaf, in cm (P3), distance between the insertion of the petiole in the leaf and the lower edge of the leaf, in cm (P2), width of the leaf, in cm (F1) and tuber weight, in kg (PT). To recognize the organization of dissimilarity of the yam genotypes, the Kohonen Self-Organizing Maps (SOM) were used. And, to perform the analysis, the Kohonen package of the RStudio software version 1.4.1103 was used. Thus, the results of the study indicated that the application of artificial intelligence through the analysis of the Kohonen Self-Organizing Map was effective for studies of genetic divergence in yams from the Recôncavo Baiano Region.

KEYWORDS: Artificial intelligence; Kohonen's Self-Organizing Maps; Genetical enhancement.

1. INTRODUÇÃO

O inhame (*Dioscorea rotundata*) pertence ao gênero *Dioscorea*, o maior e mais importante gênero da família Dioscoreaceae, apresentando mais de 600 espécies e com distribuição nas regiões tropicais, subtropicais e temperadas de todo mundo (AMANZE et al., 2011; SIQUEIRA et al., 2014). É uma planta monocotiledônea, anual e/ou perene, dioica (ATIENO et al., 2020) e propagada vegetativamente (MONDO et al., 2020), que confere grande importância social, cultural e econômica (ASIEDU; SARTIE, 2010).

Apesar da sua relevância socioeconômica e cultural, o inhame tem sofrido estresses bióticos, causados por doenças patogênicas e infestações de pragas, e abióticos, resultantes da alteração climática, condições de seca e solos pobres, acarretando a falta de material propagativo de boa qualidade, ameaçando assim a variabilidade destes recursos genéticos e ocasionando o abandono de numerosas cultivares por razões de suscetibilidade a doenças e pragas (SYOMBUA et al., 2021). Neste sentido, torna-se necessário a realização de estudos visando a variabilidade genética da espécie, uma vez que o sucesso dos programas de melhoramento depende da existência de dissimilaridade para a identificação de genótipos superiores, com intuito de desenvolver variedades melhoradas, que apresentem caracteres morfoagronômicos de interesse para os produtores e consumidores (ERTIRO et al., 2013), como resistência a doenças e pragas bem como alta produtividade com atributos apreciados pelo consumidor, conforme apontam Mignouna et al. (2007) e Obidiegwu et al. (2009), contribuindo na manutenção da sustentabilidade do agronegócio desta cultura.

Para a determinação da diversidade genética, há diversas metodologias disponíveis para estimar a dissimilaridade, seja mediante a utilização de dados fenotípicos ou com base em informações genotípicas. Entretanto, em virtude da abrangência de informações a serem analisadas e as especificidades de cada material biológico, a escolha e a correta aplicação da metodologia mais adequada são fatores importantes para a obtenção de resultados confiáveis e precisos. Assim, a inteligência artificial, como as Redes Neurais Artificiais (RNA), tem se mostrado úteis nos programas de melhoramento genético, principalmente nos estudos de detecção e reconhecimento dos padrões de organização de similaridade dos genótipos mediante a análise dos Mapas Auto-Organizáveis de

Kohonen (SOM) (BARBOSA et al., 2011; SILVA et al., 2020; CARDOSO et al., 2021).

As Redes do tipo SOM são treinadas utilizando o aprendizado competitivo e não supervisionado, em que a base dos algoritmos de adaptação dos pesos sinápticos é a competição global entre os neurônios, permitindo que a rede construa um mapa topológico a partir da apresentação dos vetores de entrada, uma vez que a rede identifica o neurônio que apresenta o vetor de peso mais semelhante ao da entrada (HAYKIN, 2001; BRAGA; CARVALHO; LUDERMIR, 2007), e assim, possibilita detectar similaridades e reconhecer padrões, com intuito de processar os agrupamentos e organizar os dados (CRUZ; NASCIMENTO, 2018).

Nesse sentido, o presente trabalho teve como objetivo estimar e organizar a divergência genética dos genótipos de inhame (*Dioscorea rotundata*) mediante a análise dos Mapas Auto-Organizáveis de Kohonen.

2. MATERIAL E MÉTODOS

2.1 Área de Estudo

Um total de 120 genótipos foram coletados em 18 propriedades rurais de povoados pertencentes aos municípios de Cruz das Almas (39° 06' 07" W, 12° 40' 12" S), São Félix (38° 58' 20" W, 12° 36' 17" S), São Felipe (39° 05' 22" W, 12° 50' 50" S) e Maragogipe (38° 55' 10" W, 12° 46' 40" S), localizados na região do Recôncavo Baiano, Brasil. O solo predominante dessas localidades é o latossolo, o clima é úmido e subúmido, com temperatura anual de 23,4°C, exceto São Felipe, que apresenta 24,4°C, e a precipitação média anual varia de 1087,7 a 1173,9 mm (SEI, 2013).

A identificação dos genótipos foi formulada por códigos com três letras, sendo que a primeira representa a cultura (I – inhame) e as subsequentes, o município no qual o genótipo foi coletado (CA - Cruz das Almas; FL – São Felipe; FX – São Félix e MA – Maragogipe, seguido pelo número do genótipo (Tabela 1).

Tabela 1. Relação dos genótipos de inhame (*Dioscorea rotundata*) coletados na Região do Recôncavo Baiano com seus respectivos códigos, série, procedência (povoado e município) e número. Cruz das Almas, BA. 2021.

Códigos dos Genótipos	Série dos Genótipos	Procedência		Nº de Genótipos
		Povoado	Município	
ICA	1 a 6	Tua 1	Cruz das Almas	6
ICA	7 a 12	Tua 2	Cruz das Almas	6
ICA	13 a 18	Cadete	Cruz das Almas	6
ICA	19 a 24	Combé	Cruz das Almas	6
ICA	25 a 30	Três Boca	Cruz das Almas	6
IFL	1 a 6	Camargo	São Felipe	6
IFL	7 a 12	Jaracandá	São Felipe	6
IFL	13 a 18	Bom Gosto	São Felipe	6
IFL	19 a 24	Boa Esperança	São Felipe	6
IFL	25 a 30	Campo das Flores	São Felipe	6
IFX	1 a 6	São Bento	São Félix	6
IFX	7 a 12	Engenho de São João	São Félix	6
IFX	13 a 18	Monte Alegre	São Félix	6
IFX	19 a 24	Boa Vista	São Félix	6
IFX	25 a 30	Matatauba	São Félix	6
IMA	1 a 6	Serraria	Maragogipe	6
IMA	7 a 12	Campinas	Maragogipe	6
IMA	13 a 18	Encruzilhada	Maragogipe	6
IMA	19 a 24	Batatan 2	Maragogipe	6
IMA	25 a 30	Batatan 3	Maragogipe	6
Total	-	-	-	120

I- Inhame; CA - Cruz das Almas; FL – São Felipe; FX – São Félix e MA – Maragogipe.

Realizou-se a caracterização dos genótipos em duas etapas. A primeira etapa ocorreu entre os meses de maio e junho de 2015, em que foram realizadas quatro visitas nos povoados dos municípios citados na Tabela 1. Em cada propriedade rural, foram selecionados aleatoriamente seis genótipos, os quais foram marcados com uma fita com a respectiva identificação de 1 a 6, sendo caracterizada a parte aérea (folhas e caules), com o auxílio de régua e paquímetro.

A segunda etapa ocorreu no mês de dezembro de 2015, em que foram realizadas duas visitas nas quais foram coletados tubérculos dos genótipos previamente marcados, ensacados e transportados para o Laboratório de Tecnologia de Alimentos da Universidade Federal do Recôncavo da Bahia (UFRB), onde foram avaliados com auxílio de fita métrica e balança.

Estes genótipos foram caracterizados com base em seis descritores quantitativos estabelecidos pelo Instituto Internacional de Recursos Genéticos Vegetais e Instituto Internacional de Agricultura Tropical (IPGRI/IITA, 1997): diâmetro do caule, em cm (DC), comprimento do pecíolo, em cm (CP), distância entre a inserção do pecíolo na folha a extremidade superior da folha, em cm (P3), distância entre a inserção do pecíolo na folha a extremidade inferior da folha, em cm (P2), largura da folha, em cm (F1) e peso do tubérculo, em kg (PT).

2.2 Redes Neurais Artificiais

Para reconhecer a organização da dissimilaridade dos genótipos de inhame, os Mapas Auto-Organizáveis de Kohonen (SOM) foram utilizados. A arquitetura da Rede Neural SOM é do tipo *feedforward* com uma camada de entrada e uma de saída, denominada mapa topológico, o qual é dividido em três etapas, segundo Cruz e Nascimento (2018), Kohonen (1990) e Kohonen (2014):

Etapa 1: Para realizar a definição do mapa topológico e o estabelecimento dos pesos aleatórios, as seis variáveis quantitativas (DC, CP, P3, P2, F1 e PT) foram utilizadas como dados de entrada da rede SOM com duas dimensões (cinco linhas e cinco colunas – 5x5) e topologia do tipo hexagonal (hextop) (Figura 1). Adotou-se o número de épocas (ou ciclos) igual a 1000, e a função de ativação utilizada para configurar as redes neurais artificiais foi a distância euclidiana.

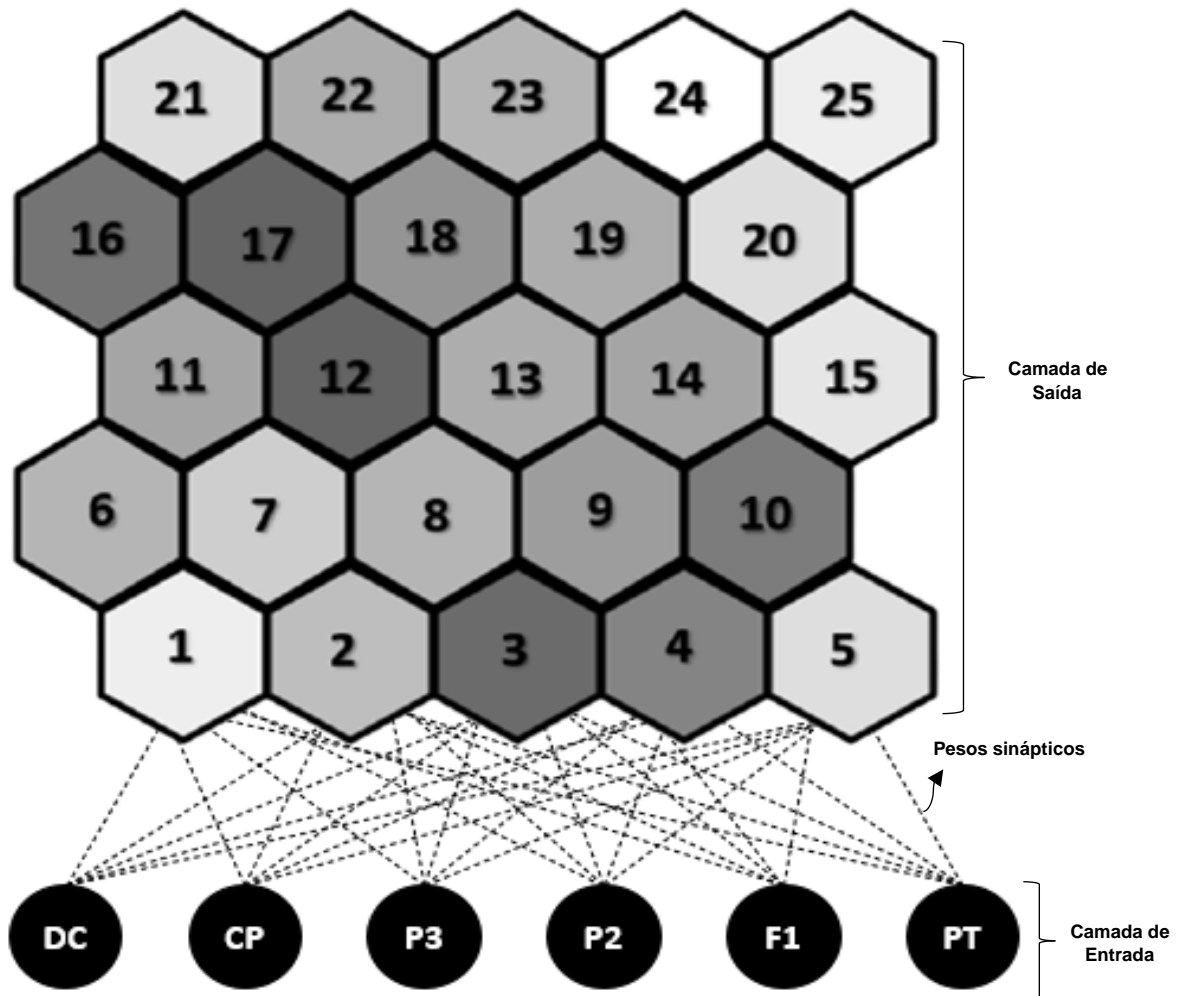


Figura 1. Arquitetura de uma rede neural SOM bidimensional com topologia do tipo hexagonal. DC: diâmetro do caule; CP: comprimento do pecíolo; P3: distância entre a inserção do pecíolo na folha a extremidade superior da folha; P2: distância entre a inserção do pecíolo na folha a extremidade inferior da folha; F1: largura da folha.

Fonte: Arquivo Pessoal (2021).

Etapa 2: Após a apresentação dos valores de entrada à rede, foi calculada a medida da distância em competição, sendo estabelecido o neurônio vencedor como o de menor distância entre eles. Ou seja, a cada padrão de entrada que é apresentado para a rede, os neurônios da camada de saída competem entre si, sendo que apenas um neurônio será o vencedor, e o neurônio vencedor é aquele que apresenta a menor distância euclidiana entre os pesos sinápticos dos neurônios e o padrão de entrada. Em seguida, os dados de entrada dos neurônios vizinhos tiveram seus pesos ajustados para determinar sua vizinhança pela taxa de aprendizado (η), sendo determinado pela seguinte expressão:

- I. $W_{(vencedor)}^{i+1} = W_{(vencedor)}^i + \eta(X_i - W_{(vencedor)}^i)$
- II. $W_{(vizinhança)}^{i+1} = W_{(vizinhança)}^i + f(x).\eta(X_i - W_{(vizinhança)}^i)$

Sendo que η é a medida da taxa de aprendizado, W é o peso dos neurônios, X_i é o vetor de entrada e $f(x)$ é a metade da taxa de aprendizado.

Etapa 3: Após todas as entradas participarem da competição, uma época (ou ciclo) foi finalizado e foi retomado a etapa 2 até que não existisse grandes diferenças entre os vetores de pesos iniciais e os ajustados.

E, para realizar a análise, utilizou-se o pacote *Kohonen* (WEHRENS; KRUISSELBRINK, 2019) do *software* RStudio versão 1.4.1103 (RSTUDIO TEAM, 2020).

3. RESULTADOS

Mediante a técnica de Redes Neurais Artificiais (RNA) do tipo SOM, foi gerado um mapa topológico (Figura 2) que consegue, a partir dos dados apresentados a rede, detectar padrões e organizá-los pelo processo de aprendizagem competitivo, uma vez que simula as conexões de neurônios e córtex cerebral por similaridades (BRAGA; CARVALHO; LUDERMIR, 2011; CRUZ; NASCIMENTO, 2018). Ou seja, esse método é capaz de ajustar os pesos sinápticos para cada variável entre as conexões de neurônios e agrupar os genótipos por padrões de caracteres comuns (CRUZ; NASCIMENTO, 2018).

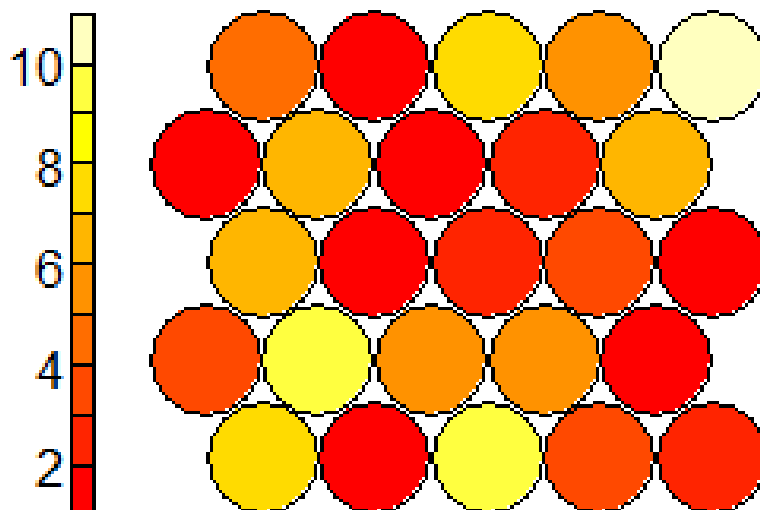


Figura 2. Organização de similaridade dos genótipos de inhame (*Dioscorea rotundata*) a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5), onde baseia-se na intensidade da coloração da área preenchida do círculo para indicar a concentração de genótipos em cada neurônio.

Fonte: Arquivo Pessoal (2021).

Pôde-se observar que os genótipos de inhame foram organizados em diferentes neurônios, representados por meio dos círculos, sendo que a intensidade do preenchimento da área dos círculos indica a concentração de genótipos contidos naquele neurônio. Assim, quanto maior a concentração de genótipos, mais clara será a área preenchida, enquanto que os neurônios que possuem uma menor quantidade de genótipos, o preenchimento do círculo é representado pela coloração mais escura (Figura 2). Então, a partir da Tabela 2, é possível analisar os genótipos que estão presentes em cada neurônio do Mapa Auto-Organizável de Kohonen, uma vez que a interpretação se dá no sentido da esquerda para a direita e de baixo para cima (Figura 2). Portanto, pôde-se perceber que não há neurônios sem elementos.

Tabela 2. Agrupamento em neurônios dos vinte povoados da Região do Recôncavo Baiano, nos quais foram cultivados os genótipos de inhame (*Dioscorea rotundata*), a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5). Cruz das Almas, BA. 2021.

NEURÔNIO	POVOADO	QUANTIDADE DE GENÓTIPOS
1	Batatan 3, Boa Esperança, Bom Gosto, Cadete, Campo das Flores, Encruzilhada, São Bento	8
2	Boa Vista	1
3	Batatan 2, Batatan 3, Boa Esperança, Cadete, Campo das Flores, Matatauba, Tua 1, Tua 2	10
4	3 Bocas, Cadete, Tua 2	4
5	Campinas, Combe, Encruzilhada	3
6	Camargo, Engenho de São João, Jacarandá, São Bento	4
7	Boa Esperança, Bom Gosto, Camargo, Combe, Jacarandá, São Bento, Serraria	10
8	Boa Vista, Cadete, Combe, Serraria	6
9	Batatan 2, Batatan 3, Campinas	6
10	Campo das Flores	1
11	3 Bocas, Batatan 3, Boa Vista, Encruzilhada, Matatauba, Tua 1	7
12	Boa Esperança	1
13	Batatan 3, Matatauba, Monte Alegre	3
14	3 Bocas, Boa Esperança, Campinas, Encruzilhada	4
15	Tua 1	1
16	Tua 2	1
17	Bom Gosto, Camargo, Jacarandá, Monte Alegre, Tua 2	7
18	Camargo	1
19	Jacarandá	3
20	3 Bocas, Batatan 2, Boa Esperança, Cadete, Camargo, Campo das Flores, Tua 1	7
21	Boa Vista, Camargo, Engenho de São João, Monte Alegre	5
22	Engenho de São João	2
23	3 Bocas, Batatan 2, Batatan 3, Cadete, Encruzilhada, Engenho de São João, São Bento	8
24	Matatauba, Monte Alegre, São Bento, Serraria	6
25	Bom Gosto, Campinas, Campo das Flores, Encruzilhada, Engenho de São João, Monte Alegre, São Bento, Serraria, Tua 1	11

Sendo assim, é importante destacar que a escolha da arquitetura da rede bem como o tamanho e a configuração do mapa fica a critério do pesquisador. Entretanto, a distribuição ideal deve ser a mais homogênea possível, ou seja, o tamanho do mapa deve ser reduzido se houver muitos neurônios vazios ou deve aumentá-lo se houver neurônios com uma densidade muito alta de genótipos (ASAN; ERCAN, 2012). Além disso, Kohonen (2001) salienta que a determinação do número de neurônios e dos parâmetros de aprendizagem é um processo empírico, mediante a experiência do profissional e em métodos de tentativa e erro. Inúmeras pesquisas estabeleceram a topologia da rede do tipo SOM recorrendo ao critério de tentativa ou de maneira aleatória (BARBOSA et al., 2011; SPANOGHE et al., 2020; WEERAKOON; SOMARATNE, 2021). Portanto, a metodologia para definir a arquitetura da rede é considerada um aspecto relevante, porque sempre que empregam redes neurais do tipo SOM nos dados coletados, os resultados obtidos podem ser diferentes, visto que as redes possuem pesos sinápticos aleatórios no início do treinamento (primeira iteração) (SANTOS et al., 2019; SILVA et al., 2020).

Além da identificação dos padrões e da organização dos genótipos, as Redes Neurais Artificiais do tipo SOM também possibilitam a distinção e classificação entre os genótipos analisados por distanciamento de neurônios. Os neurônios, onde os genótipos são alocados no mapa, são considerados regiões de classificação, onde os neurônios próximos são integrados de genótipos que possuem certa similaridade com os genótipos dos neurônios vizinhos (ASAN; ERCAN, 2012). Assim, foi gerada uma matriz de distância unificada, denominada de Matriz U, na qual as cores mais escuras indicam menor distância entre os genótipos, ou seja, são mais similares entre si, e as cores claras indicam que os genótipos são mais dissimilares (estão mais distantes) (Figura 3). Então, observou-se que a maioria dos genótipos apresentaram uma distância relevante, indicando uma presença significativa de variabilidade genética. Dessa forma, quanto maior a distância, mais se pode ter indivíduos divergentes geneticamente.

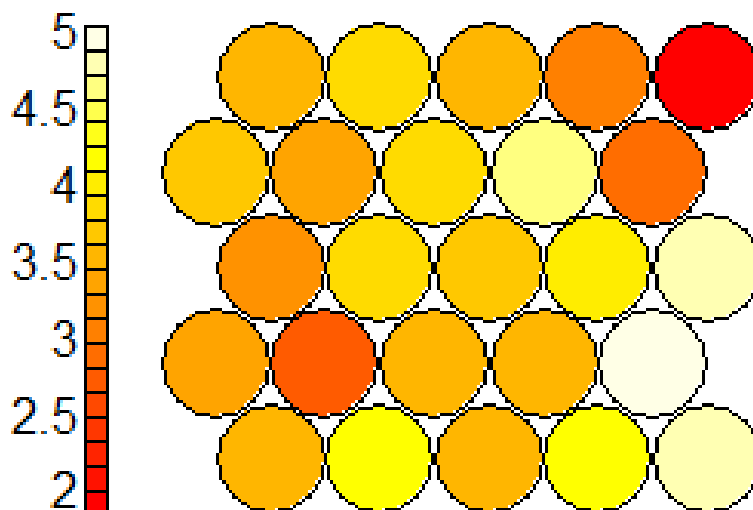


Figura 3. Matriz de distância unificada entre neurônios vizinhos dos genótipos de inhame (*Dioscorea rotundata*) a partir do Mapa Auto-Organizável de Kohonen do tipo bidimensional com configuração hexagonal (5x5).

Fonte: Arquivo Pessoal (2021).

4. DISCUSSÃO

Devido à variabilidade genética existente na cultura, o que explica sua adaptação às diferentes condições de clima e solo, possibilitando o seu cultivo em diversos países, torna-se indispensável a conservação de todo recurso disponível para evitar uma erosão genética, causada sobretudo por estresses bióticos e abióticos, fragmentação e destruição do habitat natural (por exemplo, desmatamento, desertificação, expansão urbana e modernização agrícola), a expansão de novas fronteiras agrícolas e a substituição de variedades crioulas por variedades melhoradas, as quais ocasionam a perda da diversidade genética entre e dentro das populações ao longo dos anos. Com isso, o acervo genético de uma espécie é reduzido gradualmente, limitando o seu potencial de resiliência perante essas situações e predispondo os cultivos à vulnerabilidade genética (BROWN; HODGKIN, 2015).

Os problemas ambientais sempre foram preocupações no melhoramento e, portanto, as mudanças climáticas e as adversidades geradas pelo aquecimento global têm sido agentes propulsores que estimulam e impulsionam descobertas na Ciência com o propósito de proporcionar, de forma

mais precisa, a escolha de genótipos mais produtivos e resistentes, e o desenvolvimento de meios que permitam a preservação dos recursos biológicos existentes (FONTES et al., 2002).

Então, pesquisas que visam à discriminação de genótipos dentro ou entre populações têm sido de grande relevância para o desenvolvimento de programas de melhoramento genético e para a conservação da biodiversidade. Análises da diversidade genética, mediante caracteres fenotípicos, têm contribuído efetivamente na identificação dos genótipos mais promissores, em etapas iniciais de programas de melhoramento, levando à otimização dos ganhos seletivos, devido à variabilidade encontrada nos grupos divergentes. Além disso, as análises de diversidade genética têm permitido a organização da variabilidade existente e facilitado o gerenciamento dos bancos de germoplasma, poupando tempo e recursos (BARBOSA et al., 2011).

No presente estudo, buscou-se estimar e organizar, mediante a metodologia de Mapas Auto-Organizáveis, a divergência dos genótipos de inhame. Assim, a utilização de Redes Neurais Artificiais se mostrou eficiente, uma vez que a forma como os dados foram apresentados internamente pela rede possibilitou a distinção e a organização entre os genótipos analisados por distanciamento de neurônios.

As Redes Neurais Artificiais do tipo SOM vêm se destacando como uma metodologia satisfatória e promissora para auxiliar o melhorista na organização de genótipos em programas de melhoramento, uma vez que os métodos estatísticos multivariados convencionais, como Análise de Componentes Principais (PCA), Análise Discriminante de Anderson, variáveis canônicas, métodos aglomerativos e entre outros, não são capazes de realizar essa função (ROSADO et al., 2021). Além de apresentarem uma estrutura não linear, possibilitando a sua aplicação em estudos em que as relações entre as variáveis são desconhecidas ou muito complexas (GAGO et al., 2010), demonstrando assim a capacidade de não requerer informação detalhada sobre os processos físicos do sistema a ser modelado (SUDHEER; GOSAIN; RAMASASTRI, 2003). As RNA também possibilitam a modelagem com variáveis qualitativas (ou categóricas) de forma eficiente (MARTINS et al., 2016). Além disso, as RNA não assumem nenhum tipo de distribuição dos dados a priori, diferentemente da

abordagem das técnicas tradicionais, que assumem que os dados apresentam uma distribuição normal, permitindo manipular amostras pequenas para treinamento (KAVZOGLU; MATHER, 2003) e dados faltantes (BISHOP, 1995), visto que as RNA podem ser não paramétricas (LAURINDO et al., 2017).

Atualmente, diversos estudos têm utilizado o Mapa Auto-Organizável de Kohonen para o reconhecimento do padrão de organização da diversidade genética, como Silva et al. (2020) que avaliaram o potencial da inteligência computacional em genótipos de sorgo (*Sorghum bicolor*). Rosado et al. (2021) estimaram e organizaram a dissimilaridade dos genótipos de pimenta e pimentão (*Capsicum annuum*), objetivando a seleção de cruzamentos dialélicos. Cardoso et al. (2021) analisaram a diversidade genética em genótipos de algodão fibroso colorido. Conforme observado nas pesquisas citadas, a Rede Neural Artificial do tipo SOM foi capaz de detectar as dissimilaridades e organizar as similaridades entre os acessos analisados de forma coerente e satisfatória.

A análise da dissimilaridade genética existente entre os genótipos permite informações sobre a relação de afinidade entre os mesmos, proporcionando a formação de grupos heteróticos, os quais são essenciais na escolha de genitores que possuem boa complementaridade gênica (CRUZ; REGAZZI; CARNEIRO, 2012). Portanto, a avaliação da diversidade genética é de fundamental importância em programas de melhoramento e permite a otimização da seleção de genitores, de modo que as melhores combinações híbridas sejam preditas, visto que é possível observar a variabilidade nos caracteres alvos do melhoramento genético (BARBOSA et al., 2011).

Devido os resultados promissores da utilização do Mapa Auto-Organizável de Kohonen para reconhecer o padrão de organização da dissimilaridade entre os genótipos de inhame, estudos posteriores podem ser conduzidos na perspectiva de ampliação, estabelecimento e manutenção de Bancos de Germoplasma com intuito da conservação da máxima variabilidade genética da espécie para uso imediato e futuro, visto que precisam armazenar informações que caracterizem e diferenciem os acessos, a fim de identificar genótipos que reúnam características de importância agrícola e econômica, destinados para uso em programas de melhoramento genético ou diretamente no sistema de produção da cultura.

5. CONCLUSÃO

- 1) A aplicação da inteligência artificial mediante a análise do Mapa Auto-Organizável de Kohonen foi eficaz para distinguir a dissimilaridade dos genótipos de inhame na Região do Recôncavo Baiano.
- 2) A variabilidade genética entre os genótipos de inhame pode ser explorada com o auxílio da inteligência artificial em futuros programas de melhoramento genético, uma vez que a Rede Neural do tipo SOM tem grande potencial para auxiliar o pesquisador na tomada de decisão em relação a escolha de genitores para cruzamentos dialélicos.
- 3) O entendimento da diversidade genética entre os genótipos de inhame contribui para a otimização do processo de conservação, caracterização e utilização do germoplasma, e em estudos para geração de populações segregantes e desenvolvimento de novas variedades.

REFERÊNCIAS BIBLIOGRÁFICAS

AMANZE, N. J.; AGBO, N. J.; EKE-OKORO, O. N.; NJOKU, D. N. Selection of yam seeds from open pollination for adoption in yam (*Dioscorea rotundata* Poir) production zones in Nigeria. **Journal of Plant Breeding and Crop Science**, v. 3, n. 4, p. 68-73, 2011.

ASAN, U.; ERCAN, S. An introduction to self-organizing maps. In: ZIO, E.; BAZZO, R. (org.). **Computational Intelligence Systems in Industrial Engineering**. Atlantis Press, Paris, 2012. 295-315 p.

ASIEDU, R.; SARTIE, A. Crops that feed the world 1. Yams. **Food Security**, v. 2, n. 4, p. 305-315, 2010.

ATIENO, V.; GATHERI, G. W.; KAMAU, J. W.; MUTHINI, M. Morphological and molecular characterization of cultivated yam (*Dioscorea* species) in selected counties in Kenya. **African Journal of Plant Science**, v. 14, n. 7, p. 270-279, 2020.

BARBOSA, C. D.; VIANA, A. P.; QUINTAL; S. S. R.; PEREIRA, M. G. Artificial neural network analysis of genetic diversity in *Carica papaya* L. **Crop Breeding and Applied Biotechnology**, v. 11, p. 224-231, 2011.

BISHOP, C. M. **Neural networks for pattern recognition**. Oxford: Clarendon Press, 1995. 482 p.

BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. LTC Editora, 2ª edição, 2007. 226 p.

BROWN, A. H. D.; HODGKIN, T. Indicators of genetic diversity, genetic erosion, and genetic vulnerability for plant genetic resources. In: AHUJA, M. R.; JAIN, S. M. (org.). **Genetic Diversity and Erosion in Plants**. Springer, Cham, 2015. 25-53 p.

CARDOSO, D. B. O.; MEDEIROS, L. A.; CARVALHO, G. O.; PIMENTEL, I. M.; ROJAS, G. X.; SOUSA, L. A.; SOUZA, G. M.; SOUSA, L. B. Use of computational intelligence in the genetic divergence of colored cotton plants. **Bioscience Journal**, v. 37, p. e37007-e37007, 2021.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, Minas Gerais: UFV, 2012. 508 p.

CRUZ, C. D.; NASCIMENTO, M. **Inteligência computacional aplicada ao melhoramento genético**. Viçosa, Minas Gerais: UFV, 2018. 310 p.

ERTIRO, B. T.; TWUMASI-AFRIYIE, S.; BLÜMMEL, M.; FRIESEN, D.; NEGERA, D.; WORKU, M.; ABAKEMAL, D.; KITENGE, K. Genetic variability of maize stover quality and the potential for genetic improvement of fodder value. **Field Crops Research**, v. 153, p. 79-85, 2013.

- FONTES, E. M. G.; PIRES, C. S. S.; SUJII, E. R.; PANIZZI, A. R. The environmental effects of genetically modified crops resistant to insects. **Neotropical Entomology**, v. 31, n. 4, p. 497-513, 2002.
- GAGO, J.; MATÍNEZ-NÚÑEZ, L.; LANDÍN, M.; GALLEGO, P. P. Artificial neural networks as an alternative to the traditional statistical methodology in plant research. **Journal of plant physiology**, v. 167, n. 1, p. 23-27, 2010.
- HAYKIN, S. **Redes Neurais: Princípios e Práticas**. Porto Alegre: Bookman, 2001. 900 p.
- INTERNATIONAL PLANT GENETIC RESOURCES INSTITUTE (IPGRI)/INTERNATIONAL INSTITUTE OF TROPICAL AGRICULTURE (IITA). Descriptors for yam (*Dioscorea* spp.). Roma, Itália: IPGRI/IITA, 1997. 61 p. Disponível em: <https://www.bioversityinternational.org/fileadmin/_migrated/uploads/tx_news/Descriptors_for_Yam__Dioscorea_spp.__310.pdf>. Acesso em: 15 Jun. 2021.
- KAVZOGLU, T.; MATHER, P. M. The use of backpropagating artificial neural networks in land cover classification. **International journal of remote sensing**, v. 24, n. 23, p. 4907-4938, 2003.
- KOHONEN, T. **MATLAB implementations and applications of the self-organizing map**. Helsinki, Finland: Unigrafia Oy, v. 2, 2014. 193 p.
- KOHONEN, T. **Self-organizing maps**. 3 ed. Berlin: Springer-Verlag, 2001. 502 p.
- KOHONEN, T. The self-organizing map. **Proceedings of the IEEE**, v. 78, n. 9, p. 1464-1480, 1990.
- LAURINDO, B. S.; LAURINDO, R. D. F.; AZEVEDO, A. M.; DELAZARI, F. T.; ZANUNCIO, J. C.; SILVA, J. H. Optimization of the number of evaluations for early blight disease in tomato accessions using artificial neural networks. **Scientia horticultrae**, v. 218, p. 171-176, 2017.
- MARTINS, E. R.; BINOTI, M. L. M. S.; LEITE, H. G.; BINOTI, D. H. B.; DUTRA, G. C. Configuração de redes neurais artificiais para estimação da altura total de árvores de eucalipto. **Revista brasileira de ciências agrárias**, v. 11, n. 2, p. 117-123, 2016.
- MIGNOUNA, H. D.; ABANG, M. M.; ASIEDU, R. Advances in yam (*Dioscorea* spp.) genetics and genomics. **Proceedings of the 13th ISTRC Symposium**, Arusha, p. 72-81, 2007.
- MONDO, J. M.; EDEMODOU, A.; ADEBOLA, P.; ASIEDU, R.; AKORODA, M. O.; ASFAW, A. Floral Biology and Pollination Efficiency in Yam (*Dioscorea* spp.). **Agriculture**, v. 10, n. 11, p. 560, 2020.
- OBIDIEGWU, J. E.; KOLESNIKOVA-ALLEN, M.; ENE-OBONG, E. E.; MUONEKE, C. O.; ASIEDU, R. SSR markers reveal diversity in Guinea yam (*Dioscorea cayenensis/D. rotundata*) core set. **African Journal of Biotechnology**, v. 8, n. 12, 2009.

ROSADO, R. D. S.; CECON, P. R.; OLIVEIRA, A. C. R.; FINGER, F. L.; SUELA, M. M.; CRUZ, C. D.; NASCIMENTO, M. Genetic diversity among pepper and chili genotypes by Kohonen's Self-Organizing Maps. **Genetics and Molecular Research**, v. 20, n. 3, 2021.

RSTUDIO TEAM. **RStudio: Integrated Development for R**. RStudio, PBC, Boston, MA, 2020. Disponível em: <<https://www.rstudio.com/>>.

SANTOS, I. G.; CARNEIRO, V. Q.; JUNIOR, A. C. S.; CRUZ, C. D.; SOARES, P. C. Self-organizing maps in the study of genetic diversity among irrigated rice genotypes. **Acta Scientiarum. Agronomy**, v. 41, 2018.

SILVA, M. J.; JÚNIOR, A. C. S.; CRUZ, C. D.; NASCIMENTO, M.; OLIVEIRA, M. S.; SCHAFFERT, R. E.; PARRELLA, R. A. C. Computational intelligence for studies on genetic diversity between genotypes of biomass sorghum. **Pesquisa Agropecuária Brasileira**, v. 55, 2020.

SIQUEIRA, M. V. B. M.; NASCIMENTO, W. F.; SILVA, L. R. G.; FERREIRA, A. B.; SILVA, E. F.; MING, L. C.; VEASEY, E. A. Distribution, management and diversity of yam local varieties in Brazil: a study on *Dioscorea alata* L. **Brazilian Journal of Biology**, v. 74, n. 1, p. 52-61, 2014.

SPANOGHE, M. C.; MARIQUE, T.; RIVIÈRE, J.; MOULIN, M.; DEKUIJPER, C.; NIRSHA, A.; BONNAVE, M.; LANTERBECQ, D. Genetic patterns recognition in crop species using self-organizing map: the example of the highly heterozygous autotetraploid potato (*Solanum tuberosum* L.). **Genetic Resources and Crop Evolution**, v. 67, n. 4, p. 947-966, 2020.

SUDHEER, K. P.; GOSAIN, A. K.; RAMASASTRI, K. S. Estimating actual evapotranspiration from limited climatic data using neural computing technique. **Journal of irrigation and drainage engineering**, v. 129, n. 3, p. 214-218, 2003.

SYOMBUA, E. D.; TRIPATHI, J. N.; OBIERO, G. O.; NGUU, E. K.; YANG, B.; WANG, K.; TRIPATHI, L. Potential applications of the CRISPR/Cas technology for genetic improvement of yam (*Dioscorea* spp.). **Food and Energy Security**, p. e330, 2021.

WEERAKOON, S.; SOMARATNE, S. Genetic diversity of weedy rice (*Oryza sativa* f. *spontanea*) populations in Sri Lanka: An application of Self Organizing Map (SOM). **Asian Journal of Agriculture**, v. 5, n. 1, 2021.

WEHRENS, R.; KRUISSELBRINK, J. Package 'kohonen'. 2019. Disponível em: <<https://CRAN.R-project.org/package=kohonen>>. Acesso em: 15 Out. 2021.