Model of neural networks for fertilizer recommendation and amendments in pasture crops

Rafael Hernández Moreno
Modeling and Simulation
Universidad Jorge Tadeo Lozano
Bogotá, Colombia
rafael.hernandez@utadeo.edu.co

Olmer Garcia
Engineering Department
Universidad Jorge Tadeo Lozano
Bogotá, Colombia
olmer.garciab@utadeo.edu.co

Luis Alejandro Arias R.
Department of Biological and Environmental Sciences
Universidad Jorge Tadeo Lozano
Bogotá, Colombia
luis.arias@utadeo.edu.co

***Abstract*—In the production of milk, pastures are the basis of bovine feeding, it is considered that around 50% of the costs to produce a liter of milk correspond to the feeding of the cows. Therefore the cultivation of pastures is an essential factor in the dairy industry. The cultivation of pastures requires the preparation of the soil adding fertilizers and amendments for the increase of biomass. The determination of fertilizers and amendments is an activity that adds costs to the cultivation process. The purpose of this work is to identify that from the basic nutrients in the soil such as Nitrogen (N), Phosphorus (P) and Potassium (K), with an MLP neural network of multiple input layers and multiple output layers trained with the backpropagation algorithm, it can determine the fertilizers and amendments required by pasture cultivation. The results indicate that despite using a small sample of data (44), with a threshold of 0.75 and with k-fold = 3 it was possible to identify the fertilizers Potassium chloride (Y2), Diammonium phosphate (Y6), Copper sulfate (W3) and Zinc sulfate (W4) for the cultivation of pastures. With these results, our contribution is to initiate the use of ML in pasture cultivation and later extend it to other crops. The importance of this work lies in the fact that the results can contribute to the reduction of milk production costs and that the results of this work will be the baseline for the use of other machine learning algorithms in agriculture.**

Keywords—Artificial Neural Networks, pasture crops, machine learning, data analytics.

I. Introduction

The machine learning algorithms (ML) have shown a good performance in the solution of problems of classification and prediction in real life. In communications, they have been used in sensor networks to identify network problems such as capacity, delays and routing strategies [1]. In the analysis of traffic on the Internet have been used to identify and block unwanted traffic in real time [2]. In astronomy, Support Vector Machine algorithms are used to classify galaxies based on their structure using catalogs of thousands of images supplied by the Sloan Digital Sky Survey [3]. In banking, with historical data of many clients, the ML is used to identify potential clients, offer more products and services, and analyze risks in the issuance of cards and banking products automatically [4]. ML algorithms can be classified into regression and classification the accord to the output variables. Regression models allow that from a set of attributes the result of one or more variables that are continuous can be predicted, while the classification models allow determining a class represented by quantitative values. In this work, the artificial neural networks (ANN) are used to determine the nutrients and amendments in the pasture crop. An ANN is a massively parallel processor, made up of simple processing units, which can learn from the knowledge obtained through intercommunication with other units [5]. ANN is a simplified computational model to represent the human brain [5], there is composed of processing units (neurons) and interconnections that represent the relationship between the connected nodes. The characteristics of a neural network are: Non-linearity, learning with data, adaptive and fault-tolerant [5]. Neural networks gained popularity and peaked in the early 1990s. The ideas of modern networks have not changed substantially since the 1980s, their multilayer architecture, the back-propagation algorithm, and the gradient descent are still used [6]. Nowadays, the use of ANN has been called deep learning, which is the use of neural networks with multiple layers in big data problems. So, why is it perceived as a "new" concept, if ANN has been studied since the 1940s? It is because parallel computing created by graphics processing units (GPU), distributed systems along with efficient optimization algorithms have led to the use of neural networks in contemporary/complex problems (e.g., voice recognition, search engines, and autonomous vehicles) [7].

In the following text, the first section introduces the problem overview, including an identification of the data. After that, the ANN is explained, followed by the methodology applied to obtain the results. Finally, the results and the discussion is presented showing conclusions and future studies.

II. Problem Overview

A. Pasture cultivation overview

In the world there are more than 10,000 species of grasses, the most important groups are the grasses (Poaceae) and some legumes (Fabaceae), although the latter belong to another taxonomic family are part of the grazing meadows. The grasses are monocotyledonous angiosperm plants with C_4 type photosynthesis, of narrow, simple linear leaves of striated nerves, are rich in carbohydrates that provide calories to the animals, which translates into energy to move, feed and take advantage of these foods. The legumes are dicotyledonous angiosperm plants with C_3 type photosynthesis, which have broad compound leaves of pinnated nerves and small, with accumulation of proteins that contribute to the growth and production of animals [8], in turn, are characterized by the fixation of N Atmospheric thanks to symbiotic relationships in its radical nodules, mainly with bacteria of the genus *Nostoc sp*.

Plant nutrition is the process by which a plant absorbs chemical compounds and elements of the environment that surrounds it, mainly through the roots; these are transformed into nutrients necessary for their vital metabolic processes [9]. Plants require 60 chemical elements for their growth and development [10], 16 of them are the most important for their healthy development: carbon (C), oxygen (O), nitrogen (N), phosphorus (P), potassium (K), hydrogen (H), calcium (Ca), magnesium (Mg), sulfur (S), iron (Fe), manganese (Mn), zinc (Zn), copper (Cu), boron (B), chlorine (Cl) and molybdenum (Mo). Each crop needs specific amounts of nutrients [10] that

978-1-5386-9459-6/18/$31.00 ©2018 IEEE

depend on the expected yield of the crop. These nutrients are found in chemical components, fertilizers or chemical products, Table I shows the fertilizers and Table II shows the amendments used in the cultivation of cold climate pastures [11], the latter necessary for the adaptation and/or conditioning of the ideal conditions for the absorption of nutrients by the plant, in chemical and physical terms.

TABLE I

Fertilizers

Major Fertilizers		**Minor Fertilizers**	
Product	***Variable***	***Product***	***Variable***
15 15 15 fertilizer	Y1	Iron sulfate	W1
Potassium chloride	Y2	Manganese sulfate	W2
Urea	Y3	Copper sulfate	W3
Ammonium sulfate	Y4	Zinc sulfate	W4
Calcium nitrate	Y5	Borax	W5
Diammonium phosphate	Y6		
Potassium sulfate	Y7		
Magnesium sulfate	Y8		

TABLE II

Amendments

Products	**Variable**
Organic material	Z1
Agricultural Lime (46%)	Z2
Dolomite lime	Z3
Phosphorite	Z4
Agricultural gypsum	Z5
Magnesium oxide	Z6

Among the activities of pasture crop is soil analysis, fertilizer and amendments determination [8]. Soil analysis [9] consists in identifying its characteristics such as pH, cation exchange capacity (CEC), electrical conductivity (EC), texture and the available chemical elements. The identification of fertilizers and amendments is to determine what amount of fertilizers should be added to the soil before planting and during the development of the crop, while the amendments are preferably added to the crop before the establishment of the crop to obtain optimal yield.

B. *Project location*

The data for the work is made up of 46 samples of soil analysis and their respective fertilizer recommendations and amendments for pasture cultivation in the municipality of Zipaquira, supplied by the BioSystems Center of the Jorge Tadeo Lozano University (Table III). The municipality of Zipaquira is located 43 km from Bogotá, at an altitude of 2650 meters above sea level, geographically located at 5º 01' 35" north latitude and 74º 00' 25" west longitude, with an average temperature of 11.5 °C condition that classifies it as cold climate [12] suitable for the cultivation of grasses and legumes. The surface of the municipality is 197 square kilometers [12]. In *Zipaquira* the land use distribution is 64% for livestock production, 21% for agricultural production and agricultural or livestock service activities 13% and 2% for mixed agricultural and livestock activities. The arable area for pastures is 9030 hectares.

TABLE III

Data Of Soil Analysis

Elements available in the Soil			
Major Elements			
Nitrogen (N)	X1	Iron (Fe)	X7
Phosphorus (P)	X2	Manganese (Mn)	X8
Potassium (K)	X3	Copper (Cu)	X9
Calcium (Ca)	X4	Zinc (Zn)	X10
Magnesium (Mg)	X5	Boron (B)	X11
Sulfur (S)	X6		
Soil Characteristics			
pH	X12	Electric conductivity	X13
Cation Exchange Capacity	X14	Dry weight of the sample (%)	X15
Organic material (%)	X16	Organic carbon (%)	X17
Saturation of salts			
Mg (%)	X18	Na (%)	X19
Al (%)	X20	K (%)	X21
Ca (%)	X22	Bases (%)	X23

III. Neural networks

An ANN is a massively parallel processor, consisting of simple processing units, which can learn from the knowledge obtained through intercommunication with other units [5]. Within the architectures can be distinguished the one layer perceptron [13], the multilayer perceptron (MLP) [14]and the multilayer perceptron of multiple outputs.

1) One layer perceptron: it is a processing unit, which contains an activation function. If an entry, represented by the vector $(x_1, x_{2,\dots,}, x_n)$ multiplied by their respective weights $(w_1, w_2, \dots, w_n)$ that belongs to one class (1) or another (-1), the output (o) is calculated by the activation function o $(x_1, x_{2,\dots,}, x_n)$ defined as [10]:

$$o = f(x) = \begin{cases} 1, \; if \; \sum_{i=0}^{n} w_j x_j > 0 \\ -1, \; another \; value \end{cases} \quad (1)$$

The learning process consists of determining the w_i weights according to the following rule:

$$w_i \leftarrow w_i + \Delta w_i \quad (2)$$

$$\text{Where } \Delta w_i = \gamma(t - o)x_i \quad (3)$$

Moreover, γ is learning heuristic parameter between 0 and 1.

2) Multilayer perceptron: can express a wide variety of solutions to non-linear problems, these networks seek to minimize the error between the estimated outputs (o) and the real value (t) and distribute it among all the w_{ij} weights, the training algorithm is known as Backpropagation [14]. Three types of layers are distinguished: input, hidden and output.

The general expression for the training of the hidden and output layers is found in [15] and learning consists of finding the weights according to the following rules:

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji} \quad (1)$$

Where:

w_{ji} is the associated weight from input i to unit j

- For the output layer,

$$\Delta w_{ji} = \gamma(t_j - o_j)o_j(1 - o_j)x_{ji} = \gamma\delta_j x_{ji} \quad (2)$$

$$where,\ \delta_j = (t_j - o_j)o_j(1 - o_j) \quad (3)$$

- For hidden layers,

$$\Delta w_{ji} = -\gamma\frac{\partial E_d}{\partial w_{ji}} = -\gamma\frac{\partial E_d}{\partial net_j}x_{ji} = \gamma\delta_j x_{ji} \quad (4)$$

Where, $\delta_j = o_j(1 - o_j)\sum_k \delta_k w_{kj}$

Where,

t_j is the output of unit j

o_j is the calculated output for unit j

γ is a learning constant

3) Multilayer perceptron of multiple outputs: The MIMO (Multiple Inputs with Multiple Outputs) architecture consists of an MLP where several outputs are predicted simultaneously using multiple inputs [16], [17], [18].

IV. Methodology

The problems of Machine Learning are addressed in three stages: pre-processing of the data, training, and adjustment of the model and post-processing [19]. The data for this project is found in unstructured pdf files, it was necessary to translate and transcribe the data, so before the data was pre-processed, a data transcription stage was included.

- *Pre-processing of the data.* An exploratory data analysis was performed (Table IV) and the results were visualized (Fig. 1.) in order to determine typing errors, null data, and abnormal data.

TABLE IV

Partial view of exploratory data analysis

	X1	X2	X3	X4	X5	X6
count	46	46	46	46	46	13
mean	113.04	52.83	77.53	28.41	60.39	75.46
std	39.28	48.3	43.25	34.84	50.3	54.97
min	18	0	12	0	2	6
max	150	150	150	150	150	150
mode	150	150	150	10	150	150

In this stage, the variables X6, Y1, Y3, Y7, Y8, W1, Z3, Z4, and Z5 were eliminated because they presented few values. Two observations that presented data with anomalies were eliminated. Finally, a sample of 44 observations of 22 input variables and 11 output variables was left. The null values that were presented in the variables X7, X8, X9, X19, and X11 were replaced by the mode and the average.

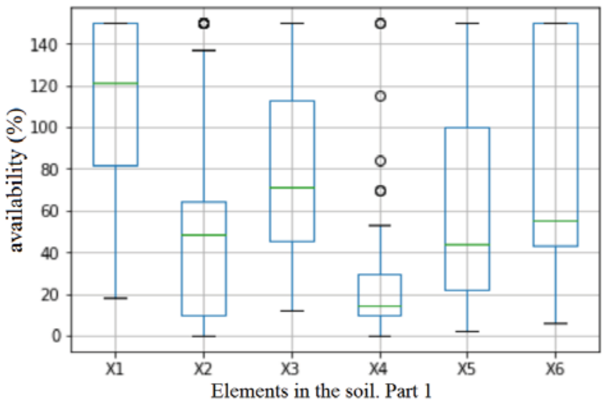

Fig.1. Partial view of the display of some variables.

The values of X and Y were normalized in the range [0,1].

- Training and adjustment of the MLP. An MLP of several outputs was chosen to determine several simultaneously y_i outputs (fertilizers and amendments recommended) that can be calculated by the linear combination of x_i inputs (nutrients available in the soil) and a group of w_j parameters that can be identify through the available data. Taking into account that a MLP of a hidden layer is able to approximate any function [20], in the algorithm for training, we used between 3 and 5 hidden layers for this purpose. For the determination of the number of neurons in each layer, the recommendation [19] was used, forming layers with a number of neurons between 11 and 100. The network was trained using the Backpropagation algorithm to minimize the mean square error (MSE) [21], for improving the results cross validation (CV) techniques was used with a k-fold of 3, 5 and 7, and:

$$MSE_i = \frac{1}{N}\sum_{J=1}^{N}\left(y_i^j - \hat{y}_i^j\right)^2 \quad (5)$$

Where y_i^j is the observation i of the fertilizer or amendment j and $\hat{y}_i^j$ is the prediction i of a fertilizer or an amendment j, for j between 1 and 11, N is the k-fold. The criteria for choosing the best model are the lowest MSE, the highest R and the highest correlation coefficient between the test observations and the estimated observations.

For the implementation, we used the scikit-learn libraries [22] of Python. These libraries have a set of APIs that allow to use the different algorithms of Machine Learning, the libraries used are MLPRegressor and GridSearchCV. MLPRegressor implements a multilayer perceptron (MLP), is trained using the Backpropagation algorithm, allows the implementation of the MIMO architecture. GridSearchCV is used to determine the number of hidden layers, number of neurons in the neural network, and the fine tune of the different components of the network using the parameters of Table III. GridSearchCV also implements the reduction of variance using cross-validation techniques, the score usde to fine tune the MLP are the mean square error, the coefficient of determination (R^2) and the correlation coefficient.

TABLE IV

Parameters used in Gridsearchcv library

Parameters	Values
Activation function	(relu, tanh)
Learning method	(lbfgs, Adam)
Iterations	(500, 1000, 1500)
Nodes in hidden layers	[(50,50,25,),(44,88,44,22), (110,88,66,22,),(100,30,80)]
Scaling	(MinMax [0,1], Standard [a,b], RobutsScaler[a,b])

V. RESULTS

A Python program was constructed using the parameters in Table IV. From the 44 observations, 36 observations (80%) were randomly chosen to train the model and 9 observations to validate it (20%). The Crossed Validation (CV) with k-fold of 3, 7 and 9 was used to decrease the variance. An MLP ANN was trained with an input layer formed by the variables of soil studies (Xi), an output layer with several simultaneous outputs (Yi) to predict fertilizers and amendments. A set of different configurations of hidden layers was dynamically configured and tested by *GridSearchCV* [22]. The program generates a summary of the network and statistics for each of the output variables (Table V). The criteria for determining the goodness of each neural network are the mean square error of the test data (Test mean) and its standard deviation (Test STD), the mean square error of the training data (Train mean) and its standard deviation (Train STD). In the Table V show the performance of neural network for different configurations of K-fold (k), activation function (AF), number of inner layer neurons and the training rule (Solver). The last three rows correspond to the models that behaved better in the prediction of the individual variables.

TABLE V

Partial view of summary statistics of many configurations of ANN

k	AF	Neurons by Layer	Solver	Test mean	Test STD	Train mean	Train STD
5	tanh	50 50 25	adam	0.290	0.010	0.192	0.019
5	tanh	100 30 80	adam	0.293	0.022	0.233	0.014
7	relu	44 88 44 22	adam	0.296	0.023	0.268	0.017
5	tanh	44 88 44 22	adam	0.297	0.019	0.268	0.019
7	relu	44 88 44 22	adam	0.299	0.020	0.220	0.018
7	relu	100 30 80	adam	0.299	0.019	0.159	0.031
3	relu	50 50 25	adam	0.303	0.028	0.113	0.003
7	relu	100 30 80	adam	0.305	0.026	0.112	0.006

In the execution of the program, the prediction of each of the 11 variables of the output layer was also performed. The correlation coefficient *Cor(Yp, Yt)* between the test data (Yt) and the predicted data (Yp) (Fig. 2.) was used as criteria to determine the quality of the model. As well as, the mean square error MSE (Yp) calculated by (5) for the test data (Yt) versus the predicted data (Yp) (Fig. 3.). The coefficient of determination R^2 of predicted data (Yp) (Fig. 4.). To choose the model. 75% was defined as a threshold for *Cor(Yp, Yt)* and R^2 (Yp). According to Figure 2, the ANN in row 7 of Table V with $k = 3$ well predicts the variables Y2, Y6, W3 and W4.

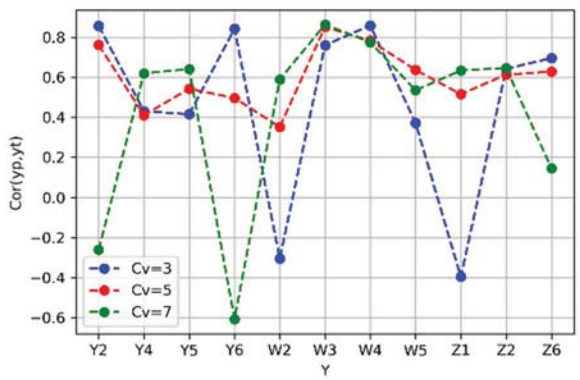

Fig. 2. Correlation between test and predicted data.

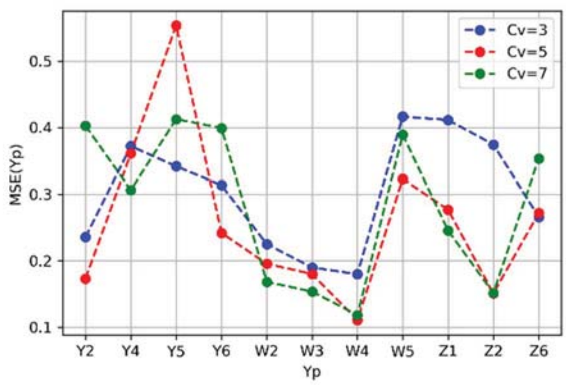

Fig. 3. Mean squared error between test and predicted data.

According to the (Fig 3), the ANN in row 7 of Table V with $k = 3$, the mean squared error trend be smaller for the variables Y2, Y6, W3 and W4.

Finally, in Fig 4 the ANN in row 7 of Table V with $k = 3$ presents a good coefficient of determination for the variables W3 and W4.

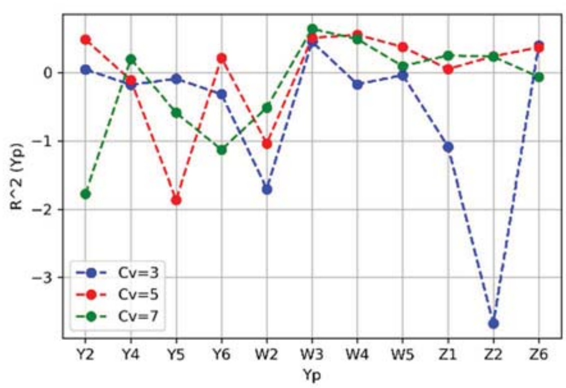

Fig. 4. Coefficient of determination between test and predicted data.

VI. CONCLUSION

There is abundant literature that indicates that Machine Learning (ML) algorithms have shown excellent performance in solving classification and prediction problems. This paper presents the application of artificial neural networks in the recommendation of fertilizers and amendments for the cultivation of cold climate pastures. As shown in figures 2 and 4, neural networks were able to predict Potassium chloride (Y2), Diammonium phosphate (Y6), Copper sulfate (W3) and Zinc sulfate (W4) with a threshold higher than 0.75.

A small number of data affected the results. The Figure 7 shows that the coefficient of determination R ^ 2 for most of the variables to predict presents high variability taking negative values, which indicates that there is not enough data [22]. Future research should include more observations to decide if the algorithm can be generalized to other regions and other types of crops. Similarly, the results showed that is possible to extend the application of ANN to other fields such as agriculture.

ACKNOWLEDGMENT

This work was possible thanks to the financial support of the Jorge Tadeo Lozano University, the BioSystems Center of the Jorge Tadeo Lozano University that provided the data set and the Modeling and Simulation Master's Degree.

REFERENCES

[1] IEEE, «A survey of machine learning in wireless sensor networks from networking and application perspectives,» de *6th International Conference on Information, Communications & Signal Processing*, Singapore, 2007.

[2] IEEE, «A survey of techniques for internet traffic classification using machine learning,» *Communications Surveys & Tutorials,* vol. 10, nº 4, 2008.

[3] SDSS, «Sloan Digital Sky Survey,» [En línea]. Available: http://www.sdss.org/surveys/. [Último acceso: 19 04 2018].

[4] C. Yeh, «The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients,» *Expert Systems with Applications,* vol. 36, nº 2, 2009.

[5] S. Haykin, Neural Networks and Learning Machines, New York: Prentice Hall., 2009.

[6] Ian Goodfellow, Y. Bengio y . C. Aaron, Deep Learning., The MIT Press., 2017.

[7] O. Garcia, C. Diaz, L. Rabelo, S. Bhide y E. Gutierrez, «Machine Learning applied to autonomous vehicles.,» de *Artificial Intelligence: Advances in Research and Applications.*, New York, Nova Science Publishers, Inc., 2018, pp. 48-74.

[8] snp power, «CULTIVO DE PASTOS. MANUAL PRÁCTICO PARA PRODUCTORES,» snp power, [En línea]. Available: http://www.swisscontact.org/fileadmin/user_upload/COUNTRIES/Peru/Documents/Publications/MANUAL_PASTOS_CULTIVADOS.pdf.

[9] D. F. Jaramillo, Introducción a la ciencia del suelo, Bogotá: Universidad Nacional, 2002.

[10] FAO, «Fertilizantes y su uso,» FAO, 2002. [En línea]. Available: http://www.fao.org/3/a-x4781s.pdf. [Último acceso: 11 05 2017].

[11] I. Gómez, «GUÍA TÉCNICA PARA EL MANEJO NUTRICINAL DE LOS CULTIVOS,» Universidad Nacional, Bogota, 2005.

[12] Oficina asesora de planeación Municipio de Zipaquirá, «Zipaquira el cambio es con todos - documento de trabajo,» 2012.

[13] F. Rosenblatt, «The perceptron: A probabilistic model for information storage and organization in the brain.,» *Psychological Review, 65(386–408).,* vol. 65, pp. 386 - 408, 1958.

[14] D. Rumelhart, «Learning Internal Representations by Error Propagation,» de *PARALLEL DISTRIBUTED PROCESSING*, vol. 323, MIT Press., 1999.

[15] T. Mitchel, Machine Learning, McGraw-Hill., 1997.

[16] A. Bezuglov, «Multi-Output Artificial Neural Network for Storm Surge Prediction in North Carolina,» Cornell University, 23 09 2016. [En línea]. Available: https://arxiv.org/abs/1609.07378.

[17] D. Antanasijević, «Multiple-input–multiple-output general regression neural networks model for the simultaneous estimation of traffic-related air pollutant emissions,» *Atmospheric Pollution Research,* vol. 9, nº 2, 2018.

[18] O. Claveria, «Multiple-input multiple-output vs. single-input single-output neural network forecasting,» Research Institute of Applied Economics, Barcelona University, Barcelona, 2015.

[19] C. Bishop, Neural Networks for Pattern Recognition., Oxford, 1999.

[20] K. Hornik, «Multilayer Feedforward Networks are Universal Approximators,» *Neural Networks,* vol. 2, 1989.

[21] T. Hastie, The Elements of Statistical Learning, Springer, 2009.

[22] F. Pedregosa, «Scikit-learn: Machine Learning in Python,» *JMLR,* nº 12, pp. 2825-2830, 2011.