Input Relevance in Multi-Layer Perceptron for Fundraising



*Published Version:*
Barro, D., Barzanti, L., Corazza, M., Nardon, M. (2024). Input Relevance in Multi-Layer Perceptron for Fundraising. Cham : Springer [10.1007/978-3-031-64273-9_6].

*Availability:*
This version is available at: https://hdl.handle.net/11585/1000929 since: 2025-01-14

*Published:*
DOI: http://doi.org/10.1007/978-3-031-64273-9_6

# Input relevance in Multi-Layer Perceptron for fundraising

Diana Barro[1], Luca Barzanti[2], Marco Corazza[1], and Martina Nardon[1]

[1] Department of Economics, Ca' Foscari University of Venice
Cannaregio 873, 30121 Venezia, Italy
d.barro@unive.it, corazza@unive.it, mnardon@unive.it
[2] Department of Mathematics, University of Bologna
Piazza di Porta San Donato 5, 40126 Bologna, Italy
luca.barzanti@unibo.it

**Abstract.** In this contribution, we consider a Multi-Layer Perceptron (MLP) methodology for predicting specific gift features, particularly the count of donations and the gift amounts. Moreover, we use Garson's indicator to evaluate the relative importance of the input variables to the output(s) in the MLP model with the aim of enhancing the effectiveness of fundraising campaigns. In the discussed application, the Donors' behaviors are estimated using a simulated dataset that includes individual characteristics and information about donation history.

**Keywords:** Multi-Layer Perceptron, Input relevance, Garson's indicator, Fundraising Management

## 1 Introduction

The optimization of a fundraising (FR) campaign, namely the maximization of the estimated global return under budget constraints, relies on the selection of the most promising Donors and requires an efficient use of available information. Accurate estimates of the number of donations, their amounts, and the gift probability are based on Donors' individual characteristics and donation history about past campaigns. These are crucial in evaluating the result of a fundraising campaign. Parametric and non-parametric approaches can be used to estimate the quantities of interest.

Recently, to estimate the gift probability [2] discusses statistical parametric methodologies and suggests modelling the number of gifts as a Poisson random variable with an intensity parameter that depends on Donors' characteristics. A Poisson regression can then be used to estimate the expected number of donations, the probability of gift, and to assign a score to each Donor measuring their propensity to the donation.

The development of non-parametric Machine Learning (ML)-based models for FR is a very recent research stream; see, for example, [4] and [3]. Along this line, [1] proposed a Multi-Layer Perceptron (MLP) to predict the number of donations and the gift amount and applied it to a simulated dataset of Donors' individual characteristics and donation history.

In this contribution, we extend the analysis carried out in [1], focusing on the relative importance of the input variables in the MLP model with the aim of enhancing the effectiveness of FR campaigns. In particular, in the analysis, we use the Garson's indicator. Section 2 presents the FR model and the dataset. Section 3 introduces the MLP and the input relevance. Section 4 discusses the application, and Section 5 concludes.

## 2 Data collection and the FR process

Associations collect in databases (DB) and manage information on Donors, Contacts and results from previous campaigns. Any gift received is associated with the Donor (a person, a company, or other entity), their available individual characteristics, and the gift history (gift events, timing and gift amounts). For large and medium-sized Associations, the information may include quantitative and qualitative features, besides advanced characteristics of the Donors' profile. For smaller ones, a systematic collection of information on Donors is very limited. The availability of data and their quantitative exploitation, together with the expertise of professionals in the field, are crucial elements in determining the success of a campaign.

In [2] and [1], the arrival of a donation, i.e. a 'gift', to an Association is modeled as the outcome of a random variable that can be analyzed according to four different dimensions: the *occurrence* of the donation, represented by a dichotomous variable; the *frequency* as a count variable measuring the number of donations received in a given period of time; the *timing* as a duration variable; and the *amount* of the donation as a positive variable.

Let $x_n$ be the vector that collects selected observable individual characteristics of Donor $n$, with $n = 1, \ldots, N$. Define $z_n$ as the vector of transformed individual characteristics, where qualitative features are properly transformed into quantitative ones or dummy variables.

Individual profile variables can be divided into personal situation variables (gender, age, number of children, education, place of origin, size of residence town, etc.); economic and financial situation variables (wage, wealth, investments); risk aversion variables (as a proxy of which the number of insurance contracts is taken); and other information such as personal interests, religious involvement, social network, geographical distances and involvement in the campaign subject, among others.

This contribution aims to assess the relevance of these input features in explaining the output elements of an FR campaign, as measured by the count of donations and the gift amounts. To this purpose, we apply and extend the MLP methodology proposed in [1].

Table 1: Some individual characteristics along the Giving Pyramid, with a finer segmentation for the Sporadic and Regular Donors

| Donors | Low wealth | Insurance policies $\geq 1$ | Min gift amount | Max gift amount |
|---|---|---|---|---|
| Sporadic (sd1) | 70 % | 35 % | 20 | 50 |
| Sporadic (sd2) | 70 % | 35 % | 30 | 100 |
| Regular (rd1) | 40 % | 65 % | 50 | 400 |
| Regular (rd2) | 40 % | 65 % | 100 | 500 |
| Large | 10 % | 65 % | 300 | 1000 |

We introduce the simulated[3] DB, which will be used to test the proposed methodology. The dataset is composed by $N = 30\,000$ Donors. The segmentation is as follows: 75 % are *Sporadic Donors* (among them, about 25 % made only one donation), 19 % *Regular Donors*, and 6 % *Large Donors*.

Personal profile variables collected are: age and number of children, education[4] (in four categories: Master and Ph.D., Bachelor, High School, other/lower school level), wealth (measured in thousands of euro), risk aversion (measured as numbers of insurance policies signed by the Donor).

The database includes the gift history for each Donor: the number of donations, the average gift amount, and the number of gift requests. Table 1 reports some individual characteristics according to the segmentation in the Giving Pyramid for the data collected in the DB.

## 3 Basics on MLP and input relevance

In this section, we first provide some basics of the supervised ML tool, known as MLP, and then we introduce the approach used to determine the relevance of the input variables to the output one(s).

According to a well-known metaphor, an MLP can be considered as a computational model inspired by the structure and functioning of the biological neural networks that comprise the brain of superior living beings. An MLP is a simple Artificial Neural Network where neurons (or nodes) represent units of computation of the network.

The nodes are organized into layers; typically: an *input layer*, whose nodes (sensors) receive the data from the external environment; one or more hidden layers, whose nodes carry out the computational tasks[5]; and an *output layer*,

[3] The DB is constructed from experts' knowledge and based on a realistic composition of a set of Donors. The database has already been used in [1].

[4] A categorical variable transformed into values ranging from 1 to 4, assigning 4 to the highest category.

[5] This represents the "intelligent" part of the computation; where the adjective intelligent means that, under mild assumptions, an MLP "*can approximate virtually any function of interest to any desired degree of accuracy* [. . .]." (see [7, p. 360].)

whose nodes (devices) release the result of the computation towards the external environment. All the nodes in one layer are fully connected to the nodes in the next one, but not among those within the same layer.

Note that in supervised ML, the MLP is trained on a labeled dataset, meaning that during the phase of parameters estimation, the MLP is presented with a dataset

$$\{(z_{1,n}, \dots, z_{i,n}, \dots, z_{I,n}; o_{1,n}, \dots, o_{k,n}, \dots, o_{K,n}), n = 1, \dots, N\},$$

where $(z_{i,n})_{i=1,\dots,I}$ is the $n$-th vector of input features, $(o_{k,n})_{k=1,\dots,K}$ is the associated vector of output labels, and $N$ is the dimension of the dataset. Pairs of nodes belonging to consecutive layers are associated with weights, namely $v_{ij}$ and $w_{jk}$, representing the strength of the connections. With reference to the MLP, the hyperparametrization phase provided the best configuration for the number of hidden layers and nodes. In our application, a single-hidden-layer MLP proved to be the best architectural structure identified for making predictions about Donor's behaviors (see [1]).

As for the training phase, the weights are adjusted in an iterative procedure over the training dataset. This stage starts with a random initialization of the weights; in subsequent runs of the MLP, the weights are chosen in order to minimize a suitable error metric based on the distance between the computed and actual outputs. The training process ends when a pre-fixed stopping criterion is satisfied, and then the obtained optimal weights are used in the validation phase.

It is well-known that the design complexity of an MLP, which can be seen as a black box, does not provide a direct interpretation of the obtained weights. In contrast, it is important to assess the impact of the explanatory inputs on the output variable(s).

Our main aim is to study the relevance of inputs to explain the outputs and to obtain relative rankings and ratings of the input features. Over the years, various methods have been proposed in the literature; for a review, see [6] and [8]. In particular, to assess the relative importance of Donors' features (input variables) on the FR campaign results (output variables), we use Garson's indicator (see [5]):

$$G_{ik} = \frac{\sum_{j=1}^{J} \left( |v_{ij} \cdot w_{jk}| \Big/ \sum_{p=1}^{I} |v_{pj}| \right)}{\sum_{q=1}^{I} \sum_{j=1}^{J} \left( |v_{qj} \cdot w_{jk}| \Big/ \sum_{p=1}^{I} |v_{pj}| \right)},$$

where $G_{ik}$ denotes the relevance of the $i$-th input variable, with $i = 1, \dots, I$, to the $k$-the output variable, with $k = 1, \dots, K$, and $J$ is the number of nodes in the hidden layer. It is noteworthy that Garson's indicator is nonnegative and normalized to sum to 1 overall input relevances, thus measuring the relative importance of the inputs but without indicating the sign of the importance itself.

## 4 Applications and results

We aim to assess the relevance of the Donors' characteristics on the count of donations and gift amounts. In our investigation, we test the following two MLP-based prediction models:

$$cd = f_{MLP,1}\left(ga, ag, nc, ed, we, ra, co\right),$$
$$ga = f_{MLP,2}\left(cd, ag, nc, ed, we, ra, co\right),$$

where $cd$ denotes the count of donations, $ga$ specifies the gift amount, $ag$, $nc$, $ed$, $we$, $ra$, and $co$ indicate the Donor's age, number of children, education level, wealth, risk aversion, and number of contacts, respectively (see Section 2). In models $f_{MLP,1}$ and $f_{MLP,2}$, the $cd$ and $ga$ forecasts are based on the other $I-1$ inputs.

On the basis of the validation results for these models, detailed in [1], we compute the related Garson's indicators.

In Tables 2a and 2b, we present the results of the analyses of input relevance for the prediction models $f_{MLP,1}$ and $f_{MLP,2}$, respectively. In both tables, in column 1, the input variables are ranked in decreasing order according to their relevance to the output, and in column 2, the values of the Garson's indicator for the same input variables are reported.

Table 2: Results related to the analyses of input relevance to the output

(a) Model $f_{MLP,1}$, with the count of donations ($cd$) as the output variable

| Input variable | $G_{i1}$ |
|---|---|
| Gift amount ($ga$) | 64.43 % |
| Risk aversion of the D. ($ra$) | 17.87 % |
| No. of children of the D. ($nc$) | 7.56 % |
| Wealth of the D. ($we$) | 5.49 % |
| Education level of the D. ($ed$) | 1.94 % |
| No. of contacts of the D. ($co$) | 1.61 % |
| Age of the D. ($ag$) | 1.11 % |

(b) Model $f_{MLP,2}$, with the gift amount ($ga$) as the output variable

| Input variable | $G_{i1}$ |
|---|---|
| Risk aversion of the D.($ra$) | 34.45 % |
| Count of donations ($cd$) | 34.34 % |
| Wealth of the D. ($we$) | 17.75 % |
| Education level of the D. ($ed$) | 6.83 % |
| No. of contacts of the D. ($co$) | 2.90 % |
| Age of the D. ($ag$) | 2.72 % |
| No. of children of the D. ($ga$) | 1.02 % |

It is worth noting that in $f_{MLP,1}$ the gift amount is the most relevant input factor to predict the count of donations; such a result is coherent with the segmentation of the Giving Pyramid and data reported in Table 1. Whereas, when considering model $f_{MLP,2}$, the count of donations is the second most relevant input in predicting the gift amount. This highlights a significant mutual dependence between these factors.

Moreover, the (decreasing) order of the input variables according to their relevance presents analogies for both prediction models. It turns out that risk aversion ($ra$) is highly relevant in both cases, and wealth ($we$) is more relevant in predicting the amount of donations, but it is also important for the number of donations. These results are also consistent with the data reported in Table 1.

Finally, in both prediction models, the first three to four most relevant input variables collectively contribute to “explain” more than 85 % to around 95 % of the associated output variable. These values are relative weights that are computed on a number $I - 1$ of available inputs. This result emphasizes the relative importance of these input variables.

## 5 Concluding remarks

In this contribution, we considered two MLP-based models to predict the number of donations and the gift amount. Among those proposed in the literature, we applied the Garson’s indicator to assess the relative importance of the input variables to the output. The main findings show that it is possible to rank the inputs based on their impact on the output and that a relatively small set of input variables can well predict the variables of interest. These results help identify which inputs are the most relevant for the forecast of the expected gift and can be exploited to implement effective FR campaigns. For future research, the analysis will be extended to consider indicators alternative to Garson’s.

## References

1. Barro, D., Barzanti, L., Corazza, M., Nardon, M.: Machine Learning and Fundraising: Applications of Artificial Neural Networks. University Ca’ Foscari of Venice, Department of Economics Research Paper Series, 33/WP/2023 (2023)
2. Barzanti, L., Nardon, M.: Estimation of the Gift Probability in Fund Raising Management. In: Corazza, M., Perna, C., Pizzi, C., Sibillo, M. (eds.): Mathematical and Statistical Methods for Actuarial Sciences and Finance. MAF 2022, 70–75, Springer, Cham (2022)
3. Cagala, T., Glogowsky, U., Rincke, J., Strittmatter, A.: Optimal targeting in Fundraising: A Machine-Learning approach. CESifo Working Papers, 9037 (2021)
4. Farrokhvar, L., Ansari, A., Kamali, B.: Predictive models for charitable giving using machine learning techniques. PLoS ONE **13(10)**, 1–14 (2021)
5. Garson, G.D.: Interpreting neural-network connection weights. Artificial Intelligence Expert, **6(4)**, 46–51 (1991)
6. Horel, E., Giesecke, K.: Significance tests for neural networks. Journal of Machine Learning Research, **21(1)**, 1–29 (2020)
7. Hornik, K., Stinchcombe, M., White, H.: Multylayer feedforward networks are universal approximators. Neural Networks **2(5)**, 359–366 (1989)
8. Mandel, F., Barnett, I.: Permutation-based hypothesis testing for neural networks. https://www.arxiv.org, arXiv:2301.11354v1 (2023)