Decisions in Economics and Finance
https://doi.org/10.1007/s10203-024-00502-9

Fundraising management through Artificial Neural Networks

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

Received: 7 October 2023 / Accepted: 28 November 2024

Abstract

In fundraising management, the assessment of the expected gift is a key point. The availability of accurate estimates of the number of donations, their amounts, and the gift probability is relevant in order to evaluate the results of a fundraising campaign. The accuracy of the expected gift estimation depends on the appropriate use of the information about Donors. In this contribution, we propose a non-parametric methodology for the prediction of Donors' behavior based on Artificial Neural Networks. In particular, Multi-Layer Perceptron is applied. In the numerical experiments, the expected gift is then estimated based on a simulated dataset of Donors' individual characteristics and information on donations history.

Keywords Fundraising management · Donor's Profile · Gift expectation · Artificial neural networks

Mathematics Subject Classification 62M45 · 62P20 · 91B74 · 91B99

JEL Classification C45 · D64

D. Barro, L. Barzanti, M. Corazza, M. Nardon: authors contributed equally to this work.

✉ Martina Nardon
mnardon@unive.it

Diana Barro
d.barro@unive.it

Luca Barzanti
luca.barzanti@unibo.it

Marco Corazza
corazza@unive.it

1 Department of Economics, University Ca' Foscari of Venice, Sestiere Cannaregio 873, 30121 Venezia, Italy

2 Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy

Published online: 13 December 2024

1 Introduction

In fundraising (FR) management, modeling the gift is crucial. The FR process can be viewed as an optimization problem: the maximization of the overall results of a campaign, subject to some restrictions and budget constraints.

The availability of accurate estimates of the gift expectation is relevant to evaluating a campaign's returns and making decisions about alternative strategies. The gift probability, the amount, number, and frequency of donations within a certain period (or for a particular campaign), and other gift features can be estimated using parametric and non-parametric approaches based on information about past donations and Donors' behavior.

Nevertheless, such information is not always available or may be very limited. In this regard, Organizations[1] can be categorized based on the existence and dimension of a structured database (DB), which may include, for each Donor, qualitative and quantitative personal profile data, in addition to the gift history. This aspect strictly depends on the Organization's size.

The success of FR strategies (the achievement of a specific FR campaign's goal and the pursuit of the Organization's mission) depends, among other factors, on the efficient use of information (see Sargeant 2001). As getting in touch with a Donor is costly, a major problem is the selection of the Contacts to maximize the expected outcome of the campaign and, at the same time, to minimize its variability. For instance, Duffy et al. (2007) deal with (potential) Donors' profiles that match some specific gift inclination to support the effectiveness of the FR process. Economists agree that information on potential Donors plays a strategic role in improving the FR results (see, for example, Nudd (2003)).

Despite the relevance of these issues, business literature and professionals in the field traditionally approached these problems with limited quantitative analysis. In the more recent literature and in the applications, we observe an evolution and a specialization of quantitative methods applied to FR management. These approaches use advanced mathematical and statistical tools, soft computing, and artificial intelligence techniques. An innovative approach has been suggested in this field by Barzanti et al. (2007) that introduces the use of mathematical modeling and Decision Support Systems (DSS) techniques. The aim is to help Associations to decide the kind of campaign to organize, the features to implement, and the Donors of the DB to contact for the maximization of the expected return of the campaign, satisfying time and budget constraints. This quantitative approach has been specialized for different types of Organizations. The contribution (Barzanti et al. 2009) considers large-sized Associations, with millions of Donors and an organizational system requiring a very sophisticated DSS. In Barzanti and Giove (2012), the focus is on small-sized Organizations and a DSS based only on essential information with no need for an organized DB. This approach has been discussed in the literature (Verhaert and Van den Poel 2012 and Melandri 2017) and validated also in the operational world by Associations (as documented in Barzanti et al. (2009) and Barzanti and Giove (2012)). In Barzanti et al. (2017), DSS targeted for medium-sized Organizations are considered. It is inter-

[1] In this work, we refer to the terms Organization and Association as synonymous.

esting to note that some similarities between the FR process and some bank activities can be set (see Moro et al. 2018); as a consequence analogous methods can be applied in the analysis.

Quantitative studies provide evidence about the main factors influencing individuals' propensity to donate. For instance, Andreoni (2006) founds that the economic and social foundations of altruism depend also on the membership to a community or the social network, and on the so-called *enlightened self-interest*. Such factors are considered by Duncan (1999) and Smith and Chang (2002). In particular, Lee et al. (1999) analyse the impact of the network of social relationships on individual's propensity to assume a role-identity as a Donor. The authors identify several factors that can impact role-identity; all these variables influence individual preferences, attitudes, and the utility people get from their decisions on how and to what extent donate (see also Cappellari et al. (2011)).

Factors that may influence the gift probability are related to individual characteristics and economic constraints: gender, age, place of origin or residence, education, number of children, financial situation, social network, personal interests, and religious involvement. Therefore, integrating all information to define an optimal FR strategy is complex.

However, tools using a classical DB approach can solve problems that are constrained by the potential of such a technology. The support to the fundraiser is limited to giving general indications in relation to specific claims without adequately managing all data about individuals. In order to improve FR strategies, experts' knowledge and advanced quantitative approaches, such as artificial intelligence, can be integrated into the process.

The analysis can be tackled at different levels. Under global perspective, one focuses on the evaluation of a campaign's overall result, while at individual level one can model the single Donor's behavior.

In this contribution, we aim at modeling those specific gift features which are relevant to evaluate the results of an FR campaign, in order to predict them as (approximate) functions of other gift features and information on Donors. To this aim, we suggest a non-parametric approach based on Machine Learning. In particular, we apply Artificial Neural Networks (ANNs) and Multi-Layer Perceptron (MLP) to predict the expected number and amounts of the donations, using as inputs some Donors' characteristics.

The remainder of this paper is organized as follows. In Sect. 2 we formally introduce the definition of the gift as an individual risk and explain how to model any aspect related to the donation. Section 3 discusses the inclusion of the individual characteristics in the Donor's profile. In Sect. 4, the numerical analyses based on ANNs are presented and compared to the results obtained from appropriate benchmark models. Finally, in Sect. 5 some concluding remarks are drawn.

2 Modeling the gift

As previously discussed, assessing an FR campaign expected return is a complex task and, to this purpose, the estimation of the expected gift is required.

The 'gift' can be modeled as an *individual risk* (see Gourieroux and Jasiak (2007)), in much analogy with other main domains of applications: finance, credit risk, insurance, and marketing.

More precisely, the gift can be viewed from four viewpoints:

- *occurrence* of a donation (the outcome is either 'yes' or 'no');
- *frequency* or *count* of donations received in a period of time (for example, a year or the duration of the campaign), so the number of gifts is zero or any positive integer;
- *timing* or *duration*, i.e. when a donation has occurred or the interval between donations,[2] whose outcome is an interval of time, usually measured with reference to a fixed point of origin (such as the beginning of the campaign or when the potential Donor has been contacted for the first time);
- *amount* of donations (the outcome is usually measured in currency units, e.g. euros, but could be also represented by working hours or other gift).

With regard to all these features, the gift is quantifiable, defining for any aspect a random variable: a dichotomous variable, a count variable, a duration variable, and a continuous positive variable, respectively.

The arrival of a donation to an Association can be treated as the outcome of a random variable, in analogy to what is done in other contexts (e.g., the arrival of a claim to an insurer, the occurrence of default in a portfolio of risks). Either dichotomous or count variables can be used to model the occurrence of the donation event. As a very simple example, consider a dichotomous random variable Y. Denote with D the gift/donation event, we have $Y = \mathbf{1}_D(\omega)$ (where $\mathbf{1}_D$ is the indicator function of D), with $\mathbb{P}[Y = 1] = p$. Then the probability of gift is equal to $\mathbb{E}(Y) = p$. Let X be a continuous random variable that represents the amount of money given by the Donor for a single donation, or the total gift of all donations filed in the considered period. In this case, the expected gift for each Donor can be computed by the product of the gift probability and expected gift amount, $\mathbb{E}(Y)\mathbb{E}(X)$.

Considering the whole campaign, both the number of gifts and the gift amount are random, hence campaign's return can be modeled as a *random sum*; in order to compute its expectation, some assumptions need to be introduced (such as independence amongst Donors, and independence of gifts count and gift amounts).

All these features can be modeled in alternative ways; however, the introduction of a realistic probability distribution may be challenging. In order to estimate the quantities of interest, both parametric and non-parametric approaches can be used, based on information about Donors and past campaigns. Recently, Barzanti and Nardon (2024) discusses statistical methodologies for modeling the gift as an individual risk, in order to estimate the gift probability. To this aim, a parametric approach has been suggested. In particular, the number of gifts is modeled as a Poisson random variable with the intensity parameter depending on Donors' individual characteristics available in the DB. The expected number of donations, and the probability of gift, can then be estimated by performing a Poisson regression, which allows also to assign a score to each Donor as a measure of their propensity to the donation.

[2] In FR management, the so-called *recency* i.e. the time length from the last donation, is particularly relevant as it is a measure of the Donor's "hotness".

3 The information on the Donor

Non-Profit Associations[3] collect and manage a variety of information to optimize their FR activity. In this process, the role of the Donor is of great importance (see, for example, Duncan (1999) and Lee et al. (1999)), as well as the choices of actions adopted by the Organization for efficiently managing the position and contacting the Donor. Practitioners claim that the $70-80\,\%$ success rate of an FR campaign derives from choosing the appropriate target of Donors to whom the strategy addresses, while only $20-30\,\%$ depends on motivations and creativity. The result of an FR campaign depends not only on Donors' profiles but also on the expertise of professionals in this field and rules of thumb.

Once a first donation is received from a Contact (i.e. a potential Donor known by the Association) or a new subject, they are labelled as 'Donor' and from that moment all the associated gift events are registered. In order to efficiently exploit the information collected on Donors and Contacts, and the experience from the past FR campaigns, *ad hoc* quantitative tools have to be developed taking into consideration the size and structure of the available dataset, and the goals to be achieved, among other features.

To describe the mechanism that gives rise to the gift, we firstly introduce some assumptions:

- any gift is associated with an individual, the Donor;
- a Donor can be a person, a company, or other entity;
- available individual characteristics of the Donor are collected in a DB;
- the gift history (gift events, timing and gift amounts) of the Donor is recorded.

For large and medium-sized Associations, the information may include both quantitative and qualitative features: information on past donations (gift history), some personal characteristics, and advanced features of the Donors' profile. Whereas small-sized Associations normally store only some quantitative information and do not use a DB to decide their strategies.

It is worth noting that using statistical methodologies, it is possible to synthesize Donor's individual characteristics with a *score* (see Barzanti and Nardon 2024). In the context of FR, such a score can be used for measuring individual propensity to donate (the higher the score, the higher the propensity to the gift), ranking Donors, and distinguishing (expected) "good" Donors. This latter procedure is called *segmentation* and can be useful to select potential Contacts or to address *ad hoc* advertising to subclasses of Donors.

Secondly, we formally define the structure of the available information in the DB. Let x_n be the vector which collects selected observable individual characteristics of Donor n, in a sample of N Donors. Define z_n as the vector of transformed individual characteristics, where qualitative features are properly transformed into quantitative ones or dummy variables.[4]

[3] With some exceptions of very small Associations.

[4] A score, summarizing the information about the Donor, can be simply defined as a scalar function of covariates $z_n'\theta$, where θ is a vector of parameters. The score can be determined using more sophisticated approaches (see Gourieroux and Jasiak 2007).

The FR literature and experts' knowledge suggest that the propensity to gift depends on some personal characteristics. Regarding the choice of personal profile variables to be used in the analysis, these can be divided into:

- personal situation variables (gender, age, number of children, education, place of origin, size of residence town);
- economic situation (wage, wealth, investments);
- risk aversion variables (the number of insurance[5] policies subscribed by the individual is taken as a proxy);
- other information (personal interests, religious involvement, social network, etc.).

Among these characteristics, the financial situation is the most significant one. Other characteristics that may have an impact are: risk aversion, geographical distance between Donor's residence and campaign location, geographical distance between Donor's interests and interests involved in the campaign, and size of residence town. The measurement of the impact of some factors can indeed be difficult, as for risk aversion. While for other factors, their influence on the gift attitude can be debatable. For instance, the presence of children can be a source of effects of opposite sign.

For most Organizations, a systematic collection of information on Donors is limited, with the exception of large Associations. Even when a DB is managed, the quantity and quality of information may be scarce. The lack of availability of data is a major drawback to the analysis. Some information cannot be collected due to different causes, depending on the instruments and the way in which donations are received (e.g. by post bulletin, rather than filling a form online), strong limitations due to the law that protects sensible data, and Donors' reluctance to provide personal information. Mistakes in the transcription or incompleteness of data, and also impossibility to assign a record to a Donor univocally identified (e.g. in case of homonymy) are causes of scarse data quality. Furthermost, managing a large DB implies for the Organizations sensible costs, expertises, efforts, and time. However, data collected in a systematized manner and efficiently used with advanced quantitative tools are major drivers to the success of the FR activity.

3.1 The data

The numerical analysis in Sect. 4 is based on a simulated DB, already used in other contributions in the literature (Barzanti et al. 2017 and Barzanti and Nardon 2024), constructed from experts' knowledge, and based on a realistic composition of a set of Donors.

The Donors' segmentation is determined by the *Giving Pyramid*, represented in Fig. 1, where the ground of the pyramid is constituted by the Contacts.

Starting with about 400 000 Contacts, a set of $N = 30\,000$ Donors is obtained. These values constitute medium to high numbers for a medium-sized Organization, or high numbers for a small-sized Organization. In the set of Donors, 75 % are *Sporadic Donors* (labeled 'sd'). Among them, about 25 % made only one donations (labeled 'sd1'), and the rest made more than one donation (labeled 'sd2'). The remaining 25 %

[5] For example, health insurance or house insurance; but also testaments are considered in this class.

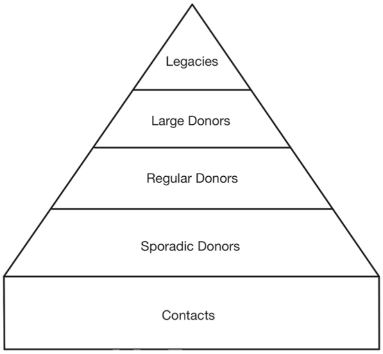

Fig. 1 Representation of the Giving Pyramid

Table 1 Distribution of some individual characteristics along the Giving Pyramid

Donors	Low wealth	Ins. policies ≥ 1	Min D amount	Max D. 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

are: 19 % *Regular Donors*[6] (labeled 'rd'), and 6 % *Large Donors*. Legacies are not present in the considered sample.

Besides information about gift history of the Donor, other personal profile variables collected are: age and number of children, education[7] (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).

Regarding the gift history, the dataset includes for each Donor: the number of donations, the gift amount for each donation,[8] and the number of gift requests (or also number of times when the Donor searched for information about the FR campaign).

Tables 1 and 2 report a synthesis of the data collected in the DB. In particular, Table 1 shows the composition (segmentation) of the Donors population in the Giving Pyramid related to some characteristics. About 70 % of the Sporadic Donors have "low wealth"; whereas, such a percentage decreases to about 40 % and 10 % for Regular Donors and Large ones, respectively. In the second column, the percentage of Donors who subscribed at least one insurance contract is reported; it can be observed that the number increases when considering higher layers of the pyramid. In the last two columns, the minimum and maximum Donation amounts are shown; in this case,

[6] A further subdivision in "stable" (labeled rd1) and "dynamic" (labeled rd2) is possible.

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

[8] The average donation is used in the analysis.

Table 2 Main statistics for the gift history and Donors' individual characteristics

	Mean	Std. Dev.	Min	Max
n. donations	6.4009	5.2036	1	28
Amount	133.6519	158.1974	20	1000
Gift requests	15.0988	8.3738	1	29
Age	53.4348	20.8576	18	89
n. children	1.4987	1.1166	0	3
Education	2.5077	1.1165	1	4
Wealth	398.4709	310.1731	10	1000
Risk aversion	1.0740	1.6726	0	5

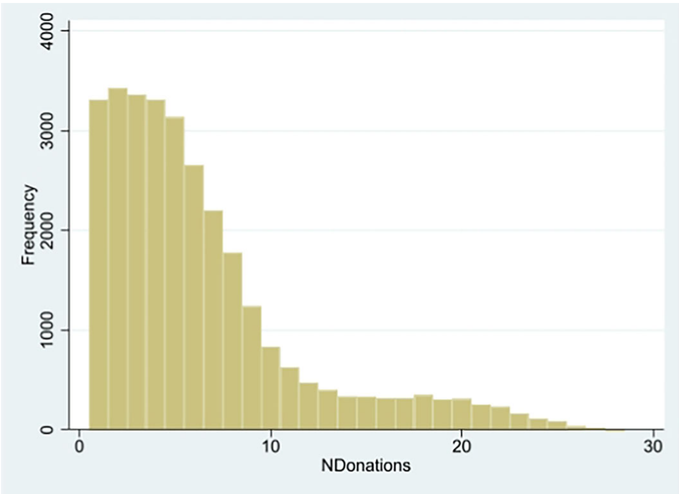

Fig. 2 Empirical distribution of the number of donations

results depend on the very definition of Sporadic (low gift amount, low frequency), Regular (low/medium gift amount, medium/high frequency), and Large (higher gift amount) Donors.

Table 2 reports the main statistics for the gift history (number of donations, amounts, number of requests), and some Donor's individual characteristics (age, number of children, education, wealth, and risk aversion).

The empirical distribution of the number of donations is shown in Fig. 2. It is worth noting that, as we considered a sample of Donors, the number of donations range from 1 to the maximum observed number. This choice allows us to avoid the inference issues associated with the excess of zeros that arise when considering all the Contacts in the DB.

4 ANNs and MLPs in FR management

In this section, we consider and apply a method for making predictions about some Donor' behaviors using a supervised Machine Learning (ML) approach known as Artificial Neural Networks (ANNs). In particular, we focus on one of the simplest ANN models, the so-called Multi Layer Perceptron (MLP).

According to a known metaphor, an ANN, and thus an MLP, can be thought of as a computational model inspired by the structure and functioning of the biological neural networks that make up the brain of the superior living beings.

In simple terms, an MLP can be viewed as a network of artificial neurons, or nodes, each of which represents a unit of computation of the network itself. These nodes are organized into layers, typically: an input layer, whose nodes receive the data from the external environment, like a sensor does; one or more hidden layers, whose nodes carry out the "intelligent" part of the computation; an output layer that releases the result of the computation towards the external environment, like a device does. By the adjective "intelligent", we mean that MLP ≪*architectures using arbitrary squashing functions can approximate virtually any function of interest to any desired degree of accuracy, provided sufficiently many hidden units are available. These results establish multilayer feedforward networks as a class of universal approximators.*≫ (see Hornik et al. 1989, p. 360). Moreover, all the nodes in one layer are fully connected to the nodes in the next layer,[9] but not among those within the same layer.

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

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

where $\left(z_{i,n}\right)_{i=1,\ldots,I}$ is the n-th vector of input features, $\left(o_{k,n}\right)_{k=1,\ldots,K}$ is the associated vector of output labels, and N is the dimension of the dataset (in the applications, N is the number of Donors).

Note also that in an MLP, pairs of nodes belonging to consecutive layers are associated with weights representing the strength of the connections. In the specific case of an MLP with one hidden layer that we consider in the application (see Sect. 4.1), pairs of nodes from the input layer to the hidden one and from the hidden layer to the output one are associated with weights v_{ji} and w_{kj}, respectively (see Fig. 3, in which an MLP with I inputs, K outputs, and one hidden layer with J nodes, is represented).

These weights are fine-tuned during the training process, based on the minimization of some error metric between the MLP's outputs and the actual outputs. In our investigation, the outputs will be the prediction of one or more Donor's behaviors (in particular, the number of donations and amount of the gift), and the errors are computed as deviations of such estimates from the past realised values of the same features collected in the DB.

[9] In the general case of an MLP with more hidden layers, the nodes in the input layer are fully connected to those in the first hidden layer, the nodes in the first hidden layer are fully connected to those in the second hidden layer, and so on, until the nodes in the last hidden layer are fully connected to those in the output layer.

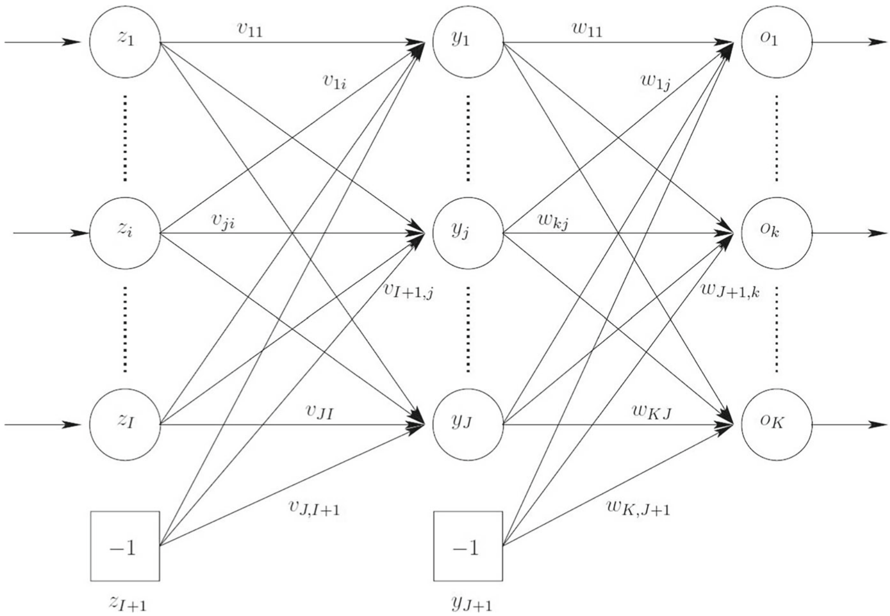

Fig. 3 Graphical representation of an MLP with one hidden layer, where: z_i, with $i = 1, \ldots, I+1$, indicate the nodes belonging to the input layer; y_j, with $j = 1, \ldots, J + 1$, denote the nodes belonging to the input layer; o_k, with $k = 1, \ldots, K$, specify the nodes belonging to the output layer; v_{ji}, with $j = 1, \ldots, J + 1$ and $i = 1, \ldots, I + 1$, indicate the weights connecting the i-th node of the input layer to the j-th node of the hidden layer; w_{kj}, with $k = 1, \ldots, K$ and $j = 1, \ldots, J + 1$, denote the weights connecting the j-th node of the hidden layer to the k-th node of the output layer. Source: Engelbrecht (2007). Note that, to avoid cluttering the figure, the squashing functions are not shown

As for the training process, it is an algorithmic procedure that adjusts in an iterative way the aforementioned weights. This process starts with a random initialization of the weights, then uses the inputs in the dataset for estimating the corresponding outputs through the MLP. The differences between the so computed outputs and the actual ones are used to appropriately update the weights in order to minimize the chosen error metric. These two steps (the output estimations, and weights updates) are repeated until a pre-fixed stopping criterion is satisfied.

In general, the training process is preceded by a more or less detailed hyperparameter tuning process. Briefly, hyperparametrization consists in appropriately setting the parameters and other features of the ANN. Once set, these parameters and features will remain fixed during the training process. For example, in the case of an MLP, this process may involve choosing the number of hidden layers, the number of nodes per hidden layer, the functional form of the squashing funtions and so on. Note that the setting of these parameters and features can heavily affect the training process, and consequently the ANN's performances.

4.1 Applications and results

The development and use of ML-based models for FR management is a very recent research area. Contributions in this field can basically be grouped in two classes: a first one in which mainly methodological proposals without or with minimal applications are presented, and a second one in which data-driven ML-based models are developed and applied.

Papers belonging to the first class include, for instance, the contribution (Philips 2022), with a discussion on how and to what extent Artificial Intelligence could be used in the FR sector. In the second class, one may cite (Farrokhvar et al. 2021), where an MLP and a Support Vector Machine are developed and applied for predicting levels of charitable giving using publicly available data sources, and Cagala et al. (2021), where Classification and Regression Decision Trees, and Classification Random Forests are used for detecting the so-called *net Donors* (that is Donors whose expected donation is higher than the marginal FR costs).

Our study fits into this second line of research. In particular, remembering that getting in touch with the Donors is costly (see Sect. 1), we aim at modeling those specific gift features which are relevant to evaluate the results of an FR campaign, namely the count of donations and the gift amounts (see Sect. 2), in order to predict them as (approximate) functions of other gift features and Donors' characteristics.

In detail, we experiment the following three MLP-based prediction models $f_{MLP,h}$, with $h = 1, 2, 3$:

- A seven-input-one-output MLP

$$(cd) = f_{MLP,1}(ga, ag, nc, ed, we, ra, gr), \tag{1}$$

  where cd denotes the count of donations, ga specifies the gift amount, ag, nc, ed, we, ra and gr indicate age, number of children, education level, wealth, risk aversion, and number of gift requests, respectively (see Sect. 3);
- A seven-input-one-output MLP

$$(ga) = f_{MLP,2}(cd, ag, nc, ed, we, ra, gr). \tag{2}$$

  This model differs from model (1) in that its output label, i.e. the gift amount ga, is one of the input features of $f_{MLP,1}$ and, vice versa, the output label of $f_{MLP,1}$, i.e. the count of donations cd, is one of the input features of $f_{MLP,2}$;
- A six-input-two-output MLP

$$(ga, cd) = f_{MLP,3}(ag, nc, ed, we, ra, gr). \tag{3}$$

  This model differs from models (1) and (2) in that it is characterized by two output labels, i.e. cd and ga, instead of one, and consequently by six input features instead of seven. It is worth noting that $f_{MLP,3}$ aims at jointly predicting both the gift features using as inputs only the Donors' characteristics.

4.1.1 The prediction model $f_{MLP,1}$

Let us first consider model $f_{MLP,1}$, defined by (1). As discussed above in this section, we initially carried out the hyperparameter tuning process, with specific reference to the number of hidden layers, and the number of nodes per hidden layer. Regarding the tuning of the other hyperparameters, we followed the suggestions of the prominent literature (see, for example, (Alpaydin 2014, sect. 11.9)).

To this end, we initially considered MLPs with 1 to 3 hidden layers and varying numbers of nodes per hidden layer. Next, we assessed three different squashing functions for the hidden nodes: the *sigmoid*, the *hyperbolic tangent*, and the *rectified linear unit* ones. As usual, we considered the linear squashing function for the output nodes. Based on the results from the error metrics, we focused on $2I + 1 = 15$ distinct MLPs with a single hidden layer, where I specifies the number of input features, having respectively from 1 to $2I + 1$ nodes in the hidden layer,[10] and we decided to use the hyperbolic tangent function.

Each of such MLPs has been trained using the dataset described in Sect. 3.1. In particular, the training phase has been performed according to the following standard steps:

- First, in order to avoid biased learning due to the order of the input–output pairs in the dataset, we shuffled the positions of these pairs.
- Then, in order to avoid overfitting in the learning phase due to an excessive MLP complexity, we used the regularization technique known as *early stopping*. This technique involves the random splitting of the original dataset in three new subdatasets, the *Training* and *Validation* ones for training purposes, and the *Testing* one for out-of-sample testing phase (for more information, see (Engelbrecht 2007, sect. 3.2)).
- Lastly, in order to manage the stochastic nature of MLP due to the random initialization of its weights, we iterated 5 times the training of each of the $2I + 1$ MLPs, and selected the best one in terms of Root Mean Square Error (RMSE) calculated over the Validation subdataset.

In Table 3, we report the results related to this first part of the hyperparameter tuning process for $f_{MLP,1}$. Observing the second column, we can detect that the minimum value of the RMSE on the Validation subdataset is reached in correspondence of an MLP having 12 nodes in the hidden layer. Therefore, in the case of $f_{MLP,1}$, the optimal number of nodes for the hidden layer is 12.

As for the second part of the hyperparameter tuning process, we experimented several configurations of MLPs with more than a single hidden layer and with different numbers of nodes in each of these layers. But none of the so configured MLPs performed better than the best one detected in the first part of the hyperparameter tuning process. Therefore, at the end of the hyperparametrization stage, the best configuration for the prediction model $f_{MLP,1}$ turned out to be a single-hidden-layer MLP with 12 nodes in the hidden layer.

Once the setting has been chosen, the prediction model has been trained using the dataset illustrated in Sect. 3.1. Furthermore, to manage the stochastic nature of MLP, we iterated 25 times this Training phase and selected the best one in terms of RMSE calculated over the Validation subdataset. In this regard, in Fig. 4, we present the

[10] Note that $2I + 1$ as upper bound for the number of nodes in the hidden layer of a single-hidden-layer MLP is a known and widely applied rule of thumb.

Table 3 Prediction model $f_{MLP,1}$. Results of the hyperparameter tuning process with respect to the number of nodes belonging to the single hidden layer

Nodes per hidden layer	RMSE on the Validation sub	Total number of weights
1	2.7716	10
2	1.9598	19
3	1.7622	28
4	1.7108	37
5	1.6560	46
6	1.6680	55
7	1.6608	64
8	1.6303	73
9	1.6509	82
10	1.6495	91
11	1.6271	100
12	1.6169	109
13	1.6250	118
14	1.6330	127
15	1.6269	136

Table 4 Prediction model $f_{MLP,1}$. Statistics related to the learning

Subdataset	RMSE	MAE	MAPE	R^2
Training	1.6491	1.1989	29.3055 %	0.8997
Validation	1.6145	1.1638	28.0060 %	0.9004
Testing	1.6200	1.1703	28.8195 %	0.9065

shapes of the RMSEs for the Training subdataset (the blue curve) and the Validation subdataset (the green curve) for one of these 25 iterations. In particular, the behavior of the RMSE over the Validation subdataset – specifically, a smooth decrease towards a (hopefully global) minimum – indicates the convergence of the learning process. Note that this specific behavior is representative of the behaviors observed in all the learning processes which have been conducted, namely the current and the subsequent ones.

In Table 4, we provide the following statistics related to the learning: RMSE, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2).

Regarding the use of R^2 as a measure of goodness of fit of ANN-based models, it has been and is under criticism for the partly unsuitableness of its applicability to nonlinear models, as MLPs are (see, for instance, Sapra (2014)). Nevertheless, R^2 continues to be widely used in the specialized literature on ML applications, even for comparative purposes among the goodness of fitting of different (nonlinear) models. In this contribution, we utilize R^2 in this latter manner.

From all the results, both the in-sample ones which are associated with the Training and Validation subdatasets, and the out-of-sample ones which are associated with the

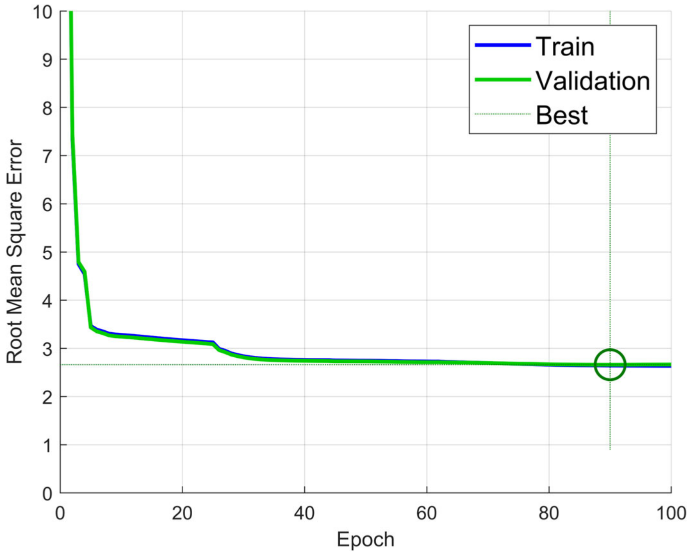

Fig. 4 Shapes of the RMSE curves for the Training and Validation subdatasets for one of the 25 iterations. In this specific iteration, the RMSE for the Validation subdataset reached its minimum at epoch 90.

Testing subdataset, we can observe that the performance of the prediction model $f_{MLP,1}$ is satisfying. In particular, we highlight that the highest value of R^2, i.e. 0.9065, has been reached in correspondence of the out-of-sample prediction.

4.1.2 The prediction model $f_{MLP,2}$

Regarding the prediction model $f_{MLP,2}$ defined in (2), we acted as for $f_{MLP,1}$. At the end of the hyperparametrization stage, the best configuration for $f_{MLP,2}$ turned out to be a single-hidden-layer MLP with 13 hidden nodes (see Table 5).

The statistics related to the learning phase are reported in Table 6. It can be observed that the performance of model $f_{MLP,2}$ in predicting the gift amount, using the count of donations and the six Donor's characteristics, is fine. The performance of model $f_{MLP,1}$ in predicting the number of donations using the gift amount and the six Donor's characteristics (as reported in Table 4) is much better. This finding highlights that the gift amount, ga, used as input for the prediction of the count of donations, cd, is more informative than cd used as input for the prediction of ga.

4.1.3 The prediction model $f_{MLP,3}$

With regard to the prediction model $f_{MLP,3}$ defined in (3), we carried out the analysis according to the steps followed for the $f_{MLP,1}$ and $f_{MLP,2}$ models. In detail, in the case of $f_{MLP,3}$, the upper bound for the number of nodes in the hidden layer is $2I + 1 = 2 \cdot 6 + 1 = 13$, since this prediction model uses as input features only the six Donors' characteristics. At the end of the hyperparametrization stage, the best

Table 5 Prediction model $f_{MLP,2}$. Results of the hyperparameter tuning process with respect to the number of nodes belonging to the single hidden layer

Nodes per hidden layer	RMSE on the Validation sub	Total number of weights
1	98.58	10
2	96.35	19
3	95.40	28
4	94.26	37
5	94.59	46
6	94.42	55
7	92.81	64
8	94.89	73
9	92.15	82
10	92.27	91
11	92.75	100
12	93.29	109
13	91.59	118
14	91.93	127
15	93.51	136

Table 6 Prediction model $f_{MLP,2}$. Statistics related to the learning

Subdataset	RMSE	MAE	MAPE	R^2
Training	91.1299	49.9592	47.2019 %	0.6698
Validation	90.0345	48.7925	45.9331 %	0.6708
Testing	89.1544	48.5719	45.9230 %	0.6842

configuration for $f_{MLP,3}$ turned out to be a single-hidden-layer MLP with 13 hidden nodes (see Table 7).

We recall that this model has two outputs (gift count and amount). In light of the fact that the model jointly predicts these two features with a reduced number of inputs, the statistics associated to the learning phase resulted poorer when compared with those of the prediction models $f_{MLP,1}$ and $f_{MLP,2}$ (see the results in Tables 8 and 4 for the count of donations, and the results in Tables 9 and 6 for the gift amount). These outcomes are expected, since in this setting two relevant sources of information are not included as explanatory variables.

However, it is worth noting that even using as inputs only the six Donors' characteristics, model $f_{MLP,3}$ shows predictive capability, in particular when considering the count of donations. In fact, the values achieved by MAPE and R^2 in this third predictive application are generally in line with those attained in a variety of other economic and financial forecasting applications (see, for instance, Zhang et al. (2019) and the references therein).

Table 7 Prediction model $f_{MLP,3}$. Results of the hyperparameter tuning process with respect to the number of nodes belonging to the single hidden layer

Nodes per hidden layer	RMSE on the Validation sub	Total number of weights
1	105.42	11
2	103.49	20
3	102.36	29
4	102.04	38
5	102.78	47
6	102.80	56
7	103.15	65
8	102.69	74
9	102.52	83
10	103.08	92
11	102.44	101
12	102.27	110
13	101.71	119

Table 8 Prediction model $f_{MLP,3}$, output labeled "count of donations" (cd). Statistics related to the learning

Subdataset	RMSE	MAE	MAPE	R^2
Training	3.9364	2.8854	64.0781 %	0.4326
Validation	3.8362	2.8044	63.9575 %	0.4274
Testing	3.9920	2.9105	65.1470 %	0.4322

Table 9 Prediction model $f_{MLP,3}$, output labeled "gift amount" (ga). Statistics related to the learning

Subdataset	RMSE	MAE	MAPE	R^2
Training	145.6249	95.4215	113.2845 %	0.1683
Validation	141.3915	92.3150	110.2419 %	0.1602
Testing	145.8508	95.8072	112.1512 %	0.1549

4.1.4 A simple benchmark check

In this section, we compare the outcomes of the prediction models $f_{MLP,1}$, $f_{MLP,2}$, and $f_{MLP,3}$ with the results from suitable linear benchmark models. The choice to use linear prediction models as benchmarks allows us to assess whether the decision to employ nonlinear predictive models, such as MLPs, has been effective.

Regarding the selection of the specific linear benchmark predictive models, we decided to use MLP models once again, but with linear activation functions for all hidden nodes. As it is known, in this case, the MLP model becomes a linear approximator of the function of interest (see, for instance, (Aggarwal 2018, sect. 1.5)). The main reasons for using such an approach instead of resorting to more conventional linear regression-based models are at least two. The first one is that MLP models do not rely on the generally stringent assumptions typically required by linear regression-based models. The second reason is that, in accordance with the principle of *Ceteris*

Table 10 Percentage variations of the learning-related statistics for each linear benchmark prediction model and for each subdataset, relative to the corresponding values for the initial nonlinear prediction models. For each error metric, i.e., RMSE, MAE, and MAPE, a positive value indicates a worsening of the metric, while a negative value indicates an improvement; the opposite holds for R^2

Prediction model, Subdataset	ΔRMSE (%)	ΔMAE (%)	ΔMAPE (%)	ΔR^2 (%)
$f_{MLP,1}$ for *cd*, Training	79.56	77.10	71.76	−42.26
$f_{MLP,1}$ for *cd*, Validation	79.18	76.82	75.54	−38.16
$f_{MLP,1}$ for *cd*, Testing	82.80	80.28	72.30	−40.86
$f_{MLP,2}$ for *ga*, Training	23.30	47.87	82.32	−25.02
$f_{MLP,2}$ for *ga*, Validation	20.87	48.53	90.32	−24.12
$f_{MLP,2}$ for *ga*, Testing	24.09	51.43	86.35	−24.90
$f_{MLP,3}$ for *cd*, Training	4.46	6.30	4.90	−36.78
$f_{MLP,3}$ for *cd*, Validation	1.18	4.79	6.76	−34.46
$f_{MLP,3}$ for *cd*, Testing	3.04	4.91	4.39	−33.70
$f_{MLP,3}$ for *ga*, Training	−0.12	−0.49	−5.51	−1.55
$f_{MLP,3}$ for *ga*, Validation	3.38	3.41	−5.08	−4.31
$f_{MLP,3}$ for *ga*, Testing	0.65	0.41	−6.56	−1.71

Paribus, by doing so we conducted the comparison while maintaining the same inferential process used for the initial nonlinear prediction models. Note that for each linear benchmark prediction models, we used the same MLP architecture and number oh hidden nodes as identified for the $f_{MLP,1}$, $f_{MLP,2}$ and $f_{MLP,3}$, respectively.

In Table 10, we present the results of the comparisons. In detail, for each linear benchmark prediction model and for each subdataset, we report the percentage variations in the values of the learning-related statistics, namely RMSE, MAE, MAPE and R^2, with respect to the corresponding values for nonlinear predictione models. Note that for each error metric, i.e., RMSE, MAE, and MAPE, a positive value indicates a worsening of the metric, while a negative value indicates an improvement; the opposite holds for R^2.

The results of the comparisons clearly show that, with a few expections, the benchmark linear prediction models perform considerably worse than $f_{MLP,1}$, $f_{MLP,2}$ and $f_{MLP,3}$. This confirms the effectiveness of the decision to employ nonlinear prediction models instead of linear ones. Only for $f_{MLP,3}$ with respect to *ga* (i.e., the gift amount), there are minor deteriorations in the metrics, along with a few slight improvements, indicating that in this case the benchmark linear approximation of the function of interest behaves almost like the nonlinear one.

5 Concluding remarks

In the organization of an FR campaign, the effective use of the information on Donors allows to optimize the resources by selecting the most promising Donors/Contacts from an organized DB for the considered context, specifying both the campaign budget and the net estimated global return. The goal is to maximize the expected global gift, under budget constraints.

The assessment of the expected gift is a crucial task, that results from the expected number of donations and gift amounts. The accuracy of these estimates depends on the efficient use of the information on Donors' individual characteristics and donations history based on past campaigns.

In this contribution, we propose the use of non-parametric models for the prediction of Donors' behavior. In particular, we applied one of the simplest ANN models, known as MLP. The obtained results indicate that these models perform particularly well (see Sect. 4.1.1) or well (see Sect. 4.1.2) if the quantities of interest are predicted separately. Furthermore, they perform satisfactorily even when these quantities are predicted jointly (see Sect. 4.1.3).

Finally, regarding future research directions, we intend to focus on the following aspects: refining the hyperparameter tuning of the MLP models to enhance their forecasting capabilities, and applying these MLP models to specific donor subclasses for tailored advertising campaigns.

Declarations

Conflict of interest The authors have no Conflict of interest (financial or non-financial) in any material discussed in this article.

References

Aggarwal, C.C.: Neural Networks and Deep Learning. A Textbook. Springer, Berlin (2018)

Alpaydin, E.: Introduction to Machine Learning, 3rd edn. MIT Press, London (2014)

Andreoni, J.: Philantropy. In: Kolm, S.C., Ythier, J. (eds.): Handbook of the Economincs of Giving, Altruism and Reciprocity, **2**, pp. 1201–1269. Elsevier (2006)

Barzanti, L., Dragoni, N., Degli Esposti, N., Gaspari, M.: Decision making in fund raising management: A knowledge based approach. In: Ellis, R., Allen, T., Petridis, M. (eds.): Applications and Innovations in Intelligent Systems, **XV**, pp. 189–201. Springer (2007)

Barzanti, L., Gaspari, M., Saletti, D.: Modelling decision making in fund raising management by a fuzzy knowledge system. Expert Syst. Appl. **36**(5), 9466–9478 (2009)

Barzanti, L., Giove, S.: A decision support system for fund raising management based on the Choquet integral methodology. Expert. Syst. **29**(4), 359–373 (2012)

Barzanti, L., Giove, S., Pezzi, A.: A Decision Support System for Non Profit Organizations. In: Petrosino, A., Loia, V., Pedrycz, W. (eds.): Fuzzy Logic and Soft Computing Applications. Lecture Notes in Artificial Intelligence, pp. 270–280. Springer (2017)

Barzanti, L., Nardon, M.: Modeling the gift in fundraising process: a parametric approach for the donations' count. Math. Meth. Econ. Fin. **17**(18), 29–44 (2024)

Cagala, T., Glogowsky, U., Rincke, J., Strittmatter, A.: Optimal targeting in Fundraising: A Machine-Learning approach. CESifo Working Papers 9037 (2021)

Cappellari, L., Ghinetti, P., Turati, G.: On time and money donations. J. Socio-Econ. **40**(6), 853–867 (2011)

Duffy, J., Ochs, J., Vesterlund, L.: Giving little by little: Dynamic voluntary contribution games. J. Public Econ. **91**(9), 1708–1730 (2007)

Duncan, B.: Modeling charitable contributions of time and money. J. Public Econ. **72**(2), 213–242 (1999)

Engelbrecht, A.P.: Computational Intelligence: An Introduction. Wiley, London (2007)

Farrokhvar, L., Ansari, A., Kamali, B.: Predictive models for charitable giving using machine learning techniques. PLoS ONE **13**(10), 1–14 (2021)

Gourieroux, C., Jasiak, J.: The Econometrics of Individual Risk. Credit, Insurance, and Marketing. Princeton University Press, Princeton (2007)

Hornik, K., Stinchcombe, M., White, H.: Multylayer feedforward networks are universal approximators. Neural Netw. **2**(5), 359–366 (1989)

Lee, L., Piliavin, J.A., Call, V.R.A.: Giving time, money, and blood: similarities and Differneces. Soc. Psycol. Quart. **62**(3), 276–290 (1999)

Melandri, V.: Fundraising. Civil Sector Press, London (2017)

Moro, S., Cortez, P., Rita, P.: A divide-and-conquer strategy using feature relevance and expert knowledge for enhancing a data mining approach to bank telemarketing. Expert. Syst. **35**(3), e12253 (2018)

Nudd, S.P.: Thinking strategically about information. In: Tempel, E. (ed.): Hank Rosso's Achieving Excellence in Fund Raising, 349–365. Wiley (2003)

Philips, V.M.: The future of Fundraising with Artificial Intelligence. ResearchGate. https://www.researchgate.net/publication/365748532 The_Future_of_Fundraising_with_Artificial_Intelligence (2022)

Sapra, R.L.: Using R^2 with caution. Curr. Med. Res. Pract. **4**(3), 130–134 (2014)

Sargeant, A.: Using donor lifetime value to inform fundraising strategy. Nonprofit Manag. Leadersh. **12**(1), 25–38 (2001)

Smith, W., Chang, C.: Shipping the good apples out: a note on contributions of time and money. Econ. Bull. **10**(1), 1–14 (2002)

Verhaert, G.A., Van den Poel, D.: The role of seed money and threshold size in optimizing fundraising campaigns: past behavior matters! Expert Syst. Appl. **39**(18), 13075–13084 (2012)

Zhang, Y., Zeng, Q., Ma, F., Shi, B.: Forecasting stock returns: do less powerful predictors help? Econ. Model. **78**(C), 32–39 (2019)