Machine Learning the Donor Journey

Anonymous Author 1 and Anonymous Author 2

[1] Institution 1, some-email@some-domain.com
[2] Institution 2, other-email@other-domain.com

Abstract. A fundamental question for charities is what action should they take next to maximize the chance of receiving a donation from a particular individual. We solve this problem using time-series data, showing that based on the previous five actions in which a charitable constituent was involved, their likely donation amount after a sixth action can be predicted. Once an accurate model is learned, the best next action can be selected to maximize donations. We show that Recurrent Neural Networks can learn accurate models of how much a constituent will donate, and we use these models to suggest actions for charities to take on an individual basis.

Keywords: machine learning, recurrent neural networks, charity

1 Introduction

In 2018, for every $100 of new charitable gifts acquired $96 was offset in losses through donor attrition [1]. Consequently, charities do not have reliable cashflow and are in a constant cycle of relying on a shrinking pool of major donors, as demographics shift. Charities seek to renew or acquire donors through *campaigns*, in which they send out an *appeal* for a specific *designation*. This communication is presently conducted primarily via email, due to the speed of delivery and the cost of direct mail communication. The number, length, and content of emails sent to constituents are generally copied from the previous year's campaign structure, or arrived at via "rules of thumb".

When charities do make data-driven decisions, they generally rely heavily on the donation history of the donor, and not on their engagement behavior, such as the actions taken by the constituent with respect to appeals sent their way. Thus, while the intention is to give each donor a personalized experience, this personalization is done on small amounts of (largely) dated donor data, ignoring what is generally the most recent donor information – the constituent's interaction with charity communications. Note that a charitable constituent is someone for whom the charity has some data (such as an email address), and who has had at least one interaction with the charity. The term "constituent" encompasses both donors and non-donors. In terms of the message being sent to these constituents, most charities use the same sequence of communications and receive similar results. Donor actions include opening an appeal email and visiting a charity donation portal.

We define the donor journey as the content and mediums of communication that a donor engages with prior to making a gift. Our goal involves modeling this journey by studying and using engagement data – how the donor interacts with the charity. The charitable world is rich with this data but there has been limited investigation into how to use machine learning techniques to prescribe actions to improve the chances of donation. Most of the data to change the sequence (and collect more money) is readily available to be collected and modelled. Businesses realize that retaining customers is more profitable than constantly acquiring new customers and take measures to reduce customer attrition [2]. But people donate for different reasons than they spend money [3], so these methods are not directly applicable to the charitable sector

In this paper, we use deep learning, specifically Recurrent Neural Networks (RNNs), to model the donor journey using the ordered actions of the constituent and charity as data, and then suggest best actions for a charity to take, on an *individual basis*. That is, for a particular constituent, what action should the charity take to increase the amount the constituent will donate. As most actions are actually performed by the constituent and out of control of the charity, in our experiments we frame this question as "what are the parameters of the next email the charity sends that will maximize expected donation?" This is a first step towards the complete personalization of charitable emails. In order to have the donor journey model trusted by real-world charities, we sought an error of $30 or lower in our models.

The next section describes related research. This is followed by a description of the problem and our approach to solve that problem. Finally, we present several empirical results using the approach and finish with discussion and future work.

2 Related Research

To the best of our knowledge, most fundraising decisions about emails are not based on any empirical evidence. Instead, anecdotal "best practices" are often shared on the web [4, 5]. In [6] the authors evaluate email campaigns from perspectives such as demographics, interest, and social network influence of their constituents, as well as external time-related factors. Basing email decisions on previous actions of the charity or the constituent has not previously been considered. In [7], direct email content is explored, through experiments with Red Cross mailings. The authors found that membership cards lead to repeat donations, while providing donors with gifts actually hurt retention. Gravyty [8] uses RNNs to *compose* emails and reinforcement learning [9] to predict the *timing* of charity actions, but do not elaborate on their approach. Our work does not focus on message composition or timing, but instead uses RNNs to choose *which* action to take with the constituent.

While charities have been hesitant to adopt machine learning techniques, for-profit organizations have long been using them, and the literature is extensive in terms of strategies for this sector [10–12]. Some of the research in machine learning for targeted marketing in for-profits can be translated to charities, but it is important to remember that these are quite distinct sectors. How people spend their money on products is not necessarily related to how they donate their money to charities. Charitable giving is not as simple (or as selfless) as one might assume [3, 13].

The use of RNNs to help predict how people will spend money, rather than donate, is common. In [14], RNNs are used to predict the customer influx and outflux at a location based on historical parking data. This is analogous to our work, as we are using the activity of constituents to predict whether they will be spending their money to help a particular charity. The work in [2] helped reduce customer attrition in e-commerce. This was done using genetic algorithms to select hyperparameters for the Artificial Neural Network actually used to classify customers, but does not make use of time-series data.

3 Problem Formulation

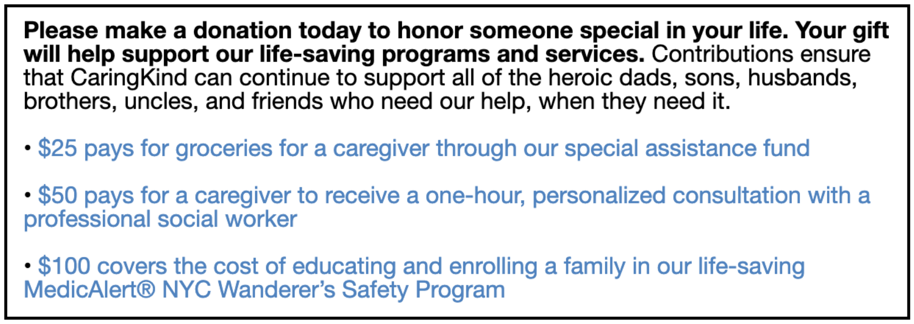
Email campaigns involve a charity sending multiple appeal emails to constituents in order to try to raise money for a designation within a charity's mandate. An example of a section of an appeal email is given in Figure 1. This message was the third in a sequence of four solicitation emails sent to constituents on the charity's targeted list. Once a constituent donated to the cause, they received no further solicitation emails, and instead received a thank you message. This demonstrates two actions that can take place during a campaign, the *delivery* of an email, and a *donation*. A full list of actions is given in Table 1.

Charities must choose which action to take next given the previous actions taken as part of a campaign, or even beyond a campaign (in the past). If a constituent has *opened* an email three times and then *clicked* a link, but not *donated*, what is the best action for the charity to take? Should the charity send an email? If so, should it be brief as to not annoy the constituent? Should it be long to provide more information to the constituent about the cause for which the charity is currently seeking funds? Should it have images to try to illustrate the cause? Should it have many links to the charity's donation portal, so a link is always on the constituents screen while they are reading the email?

All of these are valid questions encountered by fundraisers on a daily basis. In the next section, we demonstrate how RNNs can be used to automatically learn a model of a sequence of actions that leads to donation, and use the model to select actions and email parameters.

4 Our Approach

Our goal was to both model the donor journey and predict the next best action to be taken for a specific constituent. This would allow us to suggest actions and email parameters to charities that they should take or use with their constituents, to increase donations to their causes. We sought to accomplish this using RNNs.



Please make a donation today to honor someone special in your life. Your gift will help support our life-saving programs and services. Contributions ensure that CaringKind can continue to support all of the heroic dads, sons, husbands, brothers, uncles, and friends who need our help, when they need it.

• $25 pays for groceries for a caregiver through our special assistance fund

• $50 pays for a caregiver to receive a one-hour, personalized consultation with a professional social worker

• $100 covers the cost of educating and enrolling a family in our life-saving MedicAlert® NYC Wanderer's Safety Program

Fig. 1. A snapshot of an appeal email with various donation amounts to choose from.

#	**Action**	**Description**.
0	No Action	Filler action when constituent has not experienced 6 actions.
1	Delivered	An appeal emailed was successfully delivered to the constituent.
2	Opened	The constituent opened an appeal email.
3	Complained	The constituent reported an appeal email as spam.
4	Dropped	The appeal email did not reach the constituent.
5	Bounced	The appeal email was blocked by the constituent
6	Unsubscribed	The constituent unsubscribed from the charity's mailing list.
7	Pageview	The constituent viewed the donation portal.
8	Donated	The constituent made a donation.
9	Clicked	The constituent clicked on a link or donation amount.

Table 1. The list of all actions that can follow sending an email to a constituent.

RNNs using Long Short-Term Memory (LSTM) and rectified linear units (ReLU) are described in [15]. RNNs allow information to persist over time to model situations where past actions have an influence on future actions. Thus, they are an appropriate machine learning algorithm to use since we are investigating how a constituent's past actions influence their likelihood to donate to a charity in the future.

4.1 Preliminary Experimental Setup

As a first experiment in this domain, we chose to consider the previous six actions related to a constituent - both in the case where they donated, but also in the case when they did not donate. The number six was chosen since campaigns typically involve this number of constituent and charity actions. A larger window would lead to most donor journeys being padded with one or more *no action* actions, since there is no interaction between the charity and the constituent beyond these six actions. Table 1 shows the 9 actions that can follow the transmission of an email to a constituent.

An example of a series of actions that lead to a donation is {Delivered, Delivered, Pageview, Clicked, Opened, Clicked }. Here, the constituent had two emails delivered to them, then viewed the charity's donation portal. They then clicked on a suggested donation amount. Next they opened an appeal email, and clicked on the suggested amount again to donate that amount (e.g., $500).

Note that the only action in Table 1 that is controlled by the charity is "Delivered". Thus, the preliminary experiment tested not necessarily what action the charity should take next, but tested whether the RNN could learn a sequence of actions that leads to donations.

Actions were first assigned labels $\{0,1,\ldots,9\}$ with the donation amount provided last. So the previously described sequence would be $\{1,1,7,9,2,9,500\}$. The RNN predicted that actions with larger labels would lead to higher donations. One-hot encoding of the action label was used to eliminate any false sense of ordering over the actions. This was done in the standard fashion, by representing each action as a bitstring of length 10 (the number of possible actions), with all of the bits set to zero, except for the position of the action at hand.

The RNN chose only the following actions as the next best action: *pageview*, *donated*, *delivered*, *clicked* and *opened*. In 80% of cases, only the first three were chosen. This showed that the network had learned which actions lead to donations, since the first three actions are those that logically lead most directly to donations – actually donating, viewing the donation page, and the charity asking for money. The remaining two actions are also positive actions, as the constituent is engaging with the emails. Negative actions such as *bounced* and *complained* were never chosen. Having seen these results, the next step in experimentation was taken - attempting to optimize the only action a charity is capable of performing in this action set, sending an email to a constituent.

4.2 Email optimization

In email appeals the only action available to charities is actually sending an email. All other actions are taken by the constituent. We thus considered parameters of an appeal email that could be changed in order to optimize donation amount. The email parameters we considered are listed in Table 2. These were chosen as they are easy to measure, and we aim to extend this list in the future.

Parameter	**Description**.
Words	The number of words.
Paragraphs	The number of paragraphs.
Images	The number of images.
Links	The number of HTML links.
Blocks	The number of sections.
Divs	The number of HTML Content Division elements.
Variables	The number of variables to be filled.
Editable content divs	The number of editable HTML divs.

Table 2. Variable email parameters.

A sample email from Right to Play to its previous donors is shown in Figure 3. HTML divs break up an email (or any HTML document) into sections, and some of them can be edited by the charity sending out the email. The editable sections are counted in the 'editable content' parameter. Variables can be edited by the sender to allow for personalization, such as adding a constituent's name and previous donation amount. Blocks are large sections of an email that can contain multiple divs.

These parameters apply only to emails, but every action in the donor journey is either directy related to an email, or one step removed from an email – except for *no action*. Thus in addition to the one-hot encoding of the action taken, 8 additional values are added to each action, being the parameters of the associated email. If no email is associated with an action (for us, only *no action*), the 8 additional values are set to 0. As an example of an action represented in this format, consider an opened email with 235 words, 4 paragraphs, 2 images, 1 link, 2 blocks, 3 divs, 4 variables and 4 editable content sections. This action is represented as $\{0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 235, 4, 2, 1, 2, 3, 4, 4\}$. The RNN is fed a sequence of six of these actions, an example of which is shown in Figure 2. An example email is shown in Figure 3.

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	441	3	1	2	2	4	5	1
0	0	1	0	0	0	0	0	0	0	441	3	1	2	2	4	5	1
0	1	0	0	0	0	0	0	0	0	320	2	0	1	1	3	2	1
0	0	1	0	0	0	0	0	0	0	320	2	0	1	1	3	2	1
0	0	0	0	1	0	0	0	0	0	320	2	0	1	1	3	2	1

Fig. 2. A sample sequence. The cells with the darker background are the one-hot encoded actions, while the cells with the lighter background are the email parameters described in Table 2.

For almost all campaigns the set of donor data is much smaller than the set of non-donor data. Since it is true that most constituents on an email list will not give to a campaign, we experiment with feeding RNNs unbalanced training data. But, such imbalances of positive and negative data could skew any model to assume a potential donor will not donate. Thus we balance the two sets by limiting the larger (non-donor) set to be the size of the smaller (donor)

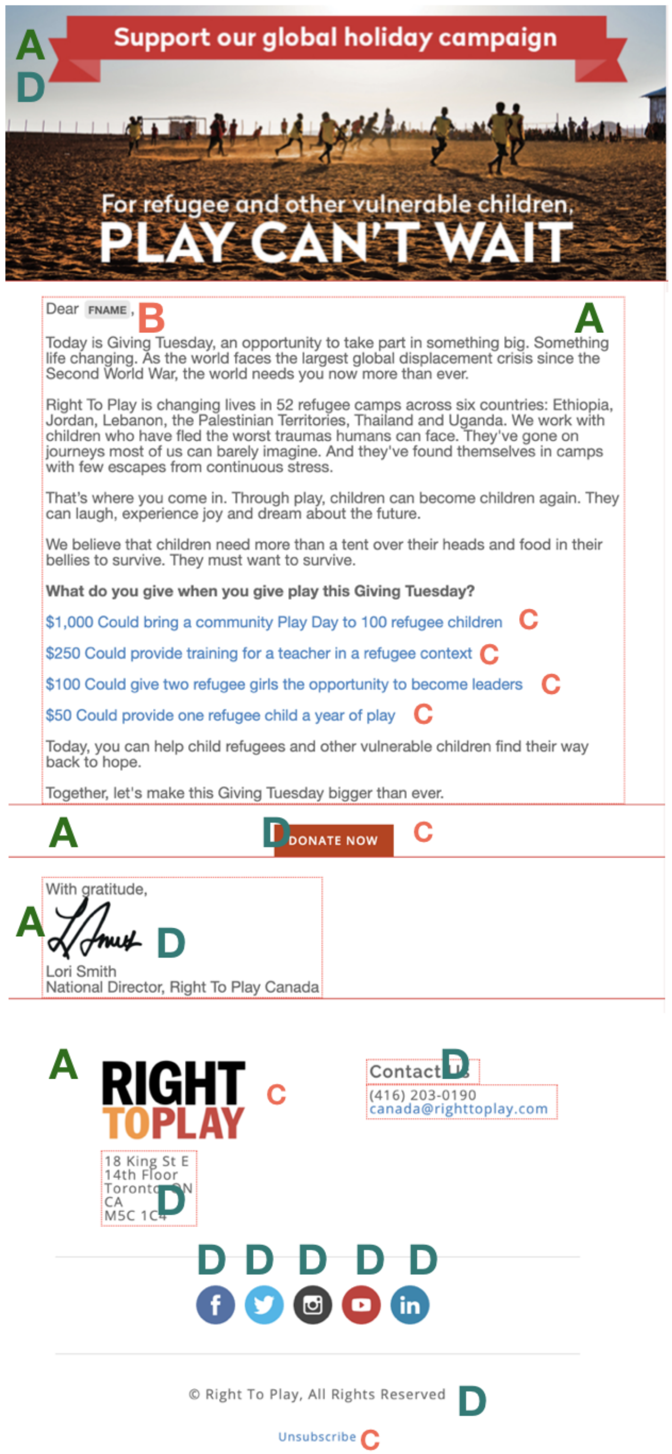


Fig. 3. A sample appeal email. There are 5 blocks (As), 1 variable (B), 7 links (Cs), 10 images (Ds), 249 words, 12 paragraphs, 48 divs (not shown), and 5 editable content divs.

set. Of course, this could lead to a poor understanding of the larger set if a misrepresentative sample were taken. Thus several iterations of this process are performed to ensure the error of this method is accurate. We experimented with both types of training data and explain this in the next section.

5 Empirical Evaluation

Given our positive preliminary experiment results (Section 4), we determined that the basic structure of the donor journey should be able to be modeled with RNNs, as the goal of this journey is to receive donations. For the eight email parameters, the values used matched to where similar to parameter values of emails sent by actual charities. Table 3 lists all values tried in various experiments. Experiments were also run with values well outside these ranges (e.g., 10000 words), but the RNN did not choose the extreme values in any case.

Parameter	**Description**
Words	{150,200, 393, 445, 457, 474, 500, 1000}
Paragraphs	{1,2,3}
Images	{2,3,6}
Links	{28,29,35}
Blocks	{8,9}
Divs	{40,41,44}
Variables	{2,3}
Editable content	{8,9,15}

Table 3. Variable email parameter values used in actual emails and experiments.

5.1 Training the RNN

Preliminary experiments were performed with data sets from three charities – a wildlife conservation charity (C1), an Alzheimer's charity (C2), and a children's charity (C3). Data was gathered from C1's Giving Tuesday appeal in 2018, C2's Spring Appeal in 2019. and C3's Year End appeal in 2018. Actions were recorded as described in Table 1. The data for C1, C2 and C3 are shown in Table 4.

For each charity, 60% of the data was used to train the RNN and 40% was used to test the learned model. Other training and testing splits tried were 75/25 and 80/20, with inferior results. This procedure was followed with both the original unbalanced sets and with balanced sets – referring to balancing the number of donors with the number of non-donors (as described in Section 4.2). As an example, using unbalanced data for C1, the RNN was trained on $(640 + 195669) * 0.6 = 117785$ constituents and tested on $(640 + 195669) * 0.4 = 78524$ constituents. For a balanced set in the same setup, the training data would be of

	C1	**C2**	**C3**
Donors	640	229	316
Non-Donors	195669	195688	50811
Total Raised	\$54,387	\$55,952	\$130,034
Minimum Donation	\$5	\$1	\$1
Maximum Donation	\$1,000	\$5,000	\$10,000

Table 4. Training data for RNN from three charities, as well as minimum and maximum donations.

size $(640 + 640) * 0.6 = 768$ constituents and the testing data was $(640 + 640) * 0.4 = 512$ constituents. Balancing the donors with the non-donors obviously dramatically reduced the training and testing data, and we ran 10 iterations to check the accuracy and error of this setup, changing the non-donors selected for training and testing in each iteration.

Various architectures for the RNN were tried for each charity, all with 18 inputs (10 actions and 8 email parameters). The best performing architectures in terms of mean absolute error (MAE) are shown in Table 5. Deeper networks did not result in more accurate predictions, nor did using a sigmoid activation function. The MAE for these architectures are shown in Table 6. For C1, The error was approximately \$25 for the balanced training data sets, which is within our margin of error to trust the model (\$30). Having accomplished our first goal, we decided to query this model as to the next best action to take with each non-donor for C1, which amounts to asking the model what should be the parameters of the email sent to each of these constituents. The error for C2 and C3 was likely higher due to the range of the donation amount in their campaigns (\$4,999 and \$9,999 vs \$995).

	C1	**C2**	**C3**
LSTM Layer 1	64	64	28
LSTM Layer 2	32	32	64
Dropout	0.2	0.5	0.5
Learning Rate	0.001	0.001	0.001
Momentum	0.9	0.9	0.9
Activation Function	ReLU	ReLU	ReLU

Table 5. Best performing RNN architectures for three charities.

5.2 Initial Email Experiments

To determine the best email parameters on a per client basis, combinations of parameter values in Table 3 were fed to a trained RNN, along with the one-hot

	C1	C2	C3
Unbalanced Training Error	$50	$236	$545
Unbalanced Testing Error	$67	$112	$493
Balanced Training Error	$29	$111	$100
Balanced Testing Error	$25	$139	$359

Table 6. MAE of predicted donations using the RNN architectures given in Table 5.

encoded *delivered* action. The best combination for non-donor was recorded – that is, the combination that led to the highest predicted donation.

The first email parameter experiments conducted varied the number of words and set all other parameter to typical values. One paragraph was used with 200 and 393 words. These two values were chosen approximately equally by the RNN as the best number of words to include in an email. Next, the number of paragraphs varied between 1 and 3, while the number of words were in the set $\{200, 393, 445, 457, 474, 500\}$, again using values from actual appeal emails used in the past. The RNN suggested 3 paragraphs with 457 words for some constituents, along with 1 paragraph with either 200 or 393 words for other constituents, showing it had learned the relationship between paragraphs and words. The final preliminary email experiment varied the number of divs $\{41, 44\}$ and links $\{28, 29\}$ along with the same sets for paragraphs and words as the previous experiment. In this experiment, the smallest number provided for each of paragraphs (1), divs (41), and links (28) were chosen the most often, with 393 words being the most chosen number of words.

5.3 Full Email Experiments

After gaining confidence in the RNN model through the preliminary experiments, a full slate of parameter values was used in order to see what type of emails are best in given situations. The parameters are the same as shown in Table 3, with the exception of *variables*, which we held at 2 after some analysis of the training data. Results were gathered on 4,261 constituents for C1, 1914 constituents for C2 and 1763 constituents for C3. While the goal of this project is to provide *personalized* email recommendations, several lessons were learned on the group as a whole. The mean, median, mode email parameter values for C1 are shown in Tables 7. Results for C2 and C3 are not shown because the model did not meet the error threshold of $30, and the model chose the same values for each constituent.

For C1, emails with fewer words (150) and paragraphs (1) were chosen most often, but for some constituents, 445 words and 3 paragraphs were selected. These constituents were those who had fewer interactions with the campaign (thus had one or more *no action* actions in their sequence). We hypothesize that the model has learned that for constituents who have received less messaging, more content in the next email will increase the donation amount. This is a first step towards our goal of personalizing email appeals.

	Mode	**Median**	**Mean**	**St. Dev**
Words	150	150	155.94	41.57
Paragraphs	1	1	1.1	0.414
Images	6	6	4.65	1.89
Links	35	29	31.7	2.99
Blocks	8	8	8.05	0.19
Divs	41	41	41	0
Variables	2	2	2	0
Editable Content	9	9	9.24	1.18

Table 7. Mean, median and mode email parameter values chosen by RNN for C1

6 Discussion and Future Work

In this paper, we showed that RNNs could learn a model of the donor journey that was within $25 of the true donation value of a constituent for a given campaign. Given that this was within our $30 threshold of business viability, we used this model to suggest actions for a charity to take next to elicit donations from non-donors, and found consistent and reasonable patterns in these predictions.

There are many extensions planned for this work. Firstly, we would like to test the email parameter selection in an actual campaign. We have agreements with some charities to do so, given the accuracy of the models. This will involve feeding current constituents into the learned RNN model and choosing the email parameters based on the model's predictions. In this way, the true performance of the predictions can be measured. The content of the messages will be determined by outside means - namely natural language processing research that is ongoing.

To strengthen the accuracy of the RNN, we will feed features of the constituents to the model, in addition to the action and email features. This will allow for more personalization, as well as give the RNN more data to learn on. These features include the constituent's demographic, donation and behaviour data. Here "behaviour data" refers to how the constituent interacts with the charity, via attending events, reading emails or watching videos. We will also experiment with journeys longer than six actions, as the extra information could help the models.

Finally, this experimental setup will be applied to potential major donors – constituents the charity is investing time into in hopes of eventually receiving large donation (typically $\$50,000+$). The set of actions taken with these constituents differs from that of a "regular" donor. We hope that we can machine learn a model to suggest the next best action to take with these constituents to increase the chances of a major gift from them. This would be another important step towards helping charities accomplish their much-needed mandates.

References

1. Levis, B., Miller, B.: 2018 Fundraising Effectiveness Survey Report. http://afpfep.org/reports/download/ (2018)

2. Fridrich, M.: Hyperparameter optimization of artificial neural network in customer churn prediction using genetic algorithm. Trends, Economics, and Management **28**(1) (2017) 9–21
3. Ptra, L., Martnez-Tur, V., Gracia, E., Moliner, C.: Why do people spend money to help vulnerable people? PLoS One **14**(3) (03 2019)
4. Jepson, T.: 11 Fundraising Email Best Practices To Drive High Response Rates, url= https://www.causevox.com/blog/11-fundraising-email-best-practices-drive-high-response-rates/, year = 2020, note = (Accessed on 01/15/2020)
5. Haydon, J.: How to write your best fundraising emails — social media today. https://www.socialmediatoday.com/content/how-write-your-best-fundraising-emails (December 2014) (Accessed on 01/16/2020).
6. Mejova, Y., Garimella, V.R.K., Weber, I., Dougal, M.: Giving is caring: understanding donation behavior through email. CSCW '14: Proceedings of the 17th ACM conference on Computer supported cooperative work and social computing (2014) 1297–1307
7. Ryzhov, I., Han, B., Bradic, J.: Cultivating disaster donors using data analytics. Management Science **62** (07 2015) 150715111657002
8. Gravyty: Why AI? Gravyty under the hood. https://www.gravyty.com/whyai/ (2020)
9. Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. MIT Press (1998)
10. Bose, I., Mahapatrab, R.: Business data mining a machine learning perspective. Information & Management **39** (2001) 211–225
11. Burez, J., Van den Poel, D.: CRM at a pay-tv company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. Expert Systems with Applications **32** (2005) 277–288
12. Apte, C., Bibelnieks, E., Natajaran, R., Pednault, E., Tipu, F., Campbell, D.: Segmentation-based modeling for advanced targeted marketing. Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining (2001) 408–413
13. Andreoni, J.: Giving with impure altruism: Applications to charity and ricardian equivalence. Journal of Political Economy **97**(6) (1989) 447–58
14. Mudassar, L., Byun, Y.C.: Customer prediction using parking logs with recurrent neural networks. International Journal of Networked and Distributed Computing **6** (2018) 133–142
15. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016) http://www.deeplearningbook.org.