This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

Voice of Charity: Prospecting the Donation Recurrence & Donor Retention in Crowdfunding

Hongke Zhao, Binbin Jin, Qi Liu, Yong Ge, Enhong Chen, *Senior Member, IEEE,* Xi Zhang, and Tong Xu

Abstract—Online donation-based crowdfunding has brought new life to charity by soliciting small monetary contributions from crowd donors to help others in trouble or with dreams. However, a crucial issue for crowdfunding platforms as well as traditional charities is the problem of high donor attrition, i.e., many donors donate only once or very few times within a rather short lifecycle and then leave. Thus, it is an urgent task to analyze the factors of and then further predict the donors behaviors. Especially, we focus on two types of behavioral events, e.g., *donation recurrence* (whether one donor will make donations at some time slices in the future) and *donor retention* (whether she will remain on the crowdfunding platform until a future time). However, this problem has not been well explored due to many domain and technical challenges, such as the *heterogeneous influence*, the *relevance of the two types of events*, and the *censoring phenomenon of retention records*. In this paper, we present a focused study on donation recurrence and donor retention with the help of large-scale behavioral data collected from crowdfunding. Specifically, we propose a Joint Deep Survival model, i.e., JDS, which can integrate heterogeneous features, e.g., donor motives, projects recently donated to, social contacts, to jointly model the donation recurrence and donor retention since these two types of behavioral events are highly relevant. In addition, we model the *censoring* phenomenon and dependence relations of different behaviors from the survival analysis view by designing multiple innovative constraints and incorporating them into the objective functions. Finally, we conduct extensive analysis and validation experiments with large-scale data collected from Kiva.org. The experimental results clearly demonstrate the effectiveness of our proposed models for analyzing and predicting the donation recurrence and donor retention in crowdfunding.

Index Terms—Crowdfunding, Donor Retention, Survival Analysis, Ranking Constraints, Deep Learning.

1 Introduction

Crowdfunding is an emerging Internet-based fundraising mechanism soliciting small monetary contributions from crowd donors to help others in trouble or with dreams [1]. Recent years have witnessed the rapid development of crowdfunding platforms among which the donation-based ones are becoming increasingly popular [1], [2], such as Kiva.org [1] [3], and DonorsChoose.org [2] [4]. Leveraging Internet, crowdfunding has brought new life to charity, i.e., making it easy to donate any amount of money even every penny to help others across the globe.

For example, Kiva.org is an international nonprofit platform, founded in 2005, with a mission to connect people through lending to alleviate poverty. Specifically, Kiva.org enables *Field Partners* (nonprofit organizations around the world) to screen the needy or suffering, and post requests in the form of projects to Kiva.org for funding. Then the accessing *donors* crowdfund these projects in increments of $25 or more. Donors may act as individuals or *teams*. The critical component for the success of crowdfunding communities is the recruitment and continued engagement of donors [4]. However, because of the non-profit nature, the situation relating to donor retention for donation-based crowdfunding as well as traditional charities is extremely serious, i.e., usually, the donor attrition rate is above 70% [4]. Actually, customer attrition/churn [5], [6] is crucial and highly focused on in many commercial scenarios, such as E-commerce, finance and services. However, for a quite long time, relevant studies on analyzing donor retention in charity have been rather limited in the literature.

Fortunately, with the accumulation of large-scale user behavior data in crowdfunding platforms, many data-driven studies which focus on analyzing the user behaviors have been conducted [7], [8]. For example, Liu, et al. [7] studied the donation motivation classification in Kiva.com. Especially, Althoff, et al. [4] explored various factors impacting donor retention in DonorsChoose.org from the statistical perspectives which was inspiring for our research. However, how to comprehensively analyze the heterogeneous factors affecting and then further predict the *donor retention or attrition*, are still largely unexplored areas, both in the charity and in other domains. In addition to these heterogeneous factors, according to our observation and analysis, donors' own behaviors (i.e., *donation recurrence*) could particularly reflect their decision on retention. In fact, donation recurrence prediction is an inevitable intermediate goal for predicting the donor retention. Thus, in this paper, we attempt to track this problem by jointly predicting the donor retention and also the intermediate goal (predicting donation recurrence). Although it is necessary to construct the predictions of donation recurrence and donor retention,

- *H. Zhao, B. Jin, Q. Liu, E. Chen and T. Xu are with Anhui Province Key Laboratory of Big Data Analysis and Application, School of Computer Science and Technology, University of Science and Technology of China, Hefei, Anhui 230026, China. E. Chen is the corresponding author. E-mail:{zhhk,bb0725}@mail.ustc.edu.cn, {qiliuql,cheneh,tongxu}@ustc.edu.cn.*
- *Y. Ge is with the Department of Management Information Systems, Eller College of Management, University of Arizona, Tucson, AZ 85721. E-mail: yongge@email.arizona.edu.*
- *X. Zhang is with the College of Management and Economics, Tianjin University, Tianjin 300072, China. E-mail: jackyzhang@tju.edu.cn.*

Manuscript received July 20, 2018.

1. https://www.kiva.org/
2. https://www.donorschoose.org/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

as they can alert platforms that they need to do something before they lose donors, this is a very challenging task.

First, donor behaviors, e.g., donation recurrence, donor retention or attrition, are influenced by various factors [4], such as their motives and preferences, their social contacts in crowdfunding communities, and the characteristics of the projects to which they have recently donated. How to comprehensively analyze the heterogeneous features and integrate them for accurate prediction is not a trivial issue. Second, according to our data analysis, the behaviors of donors, especially the donation recurrence, are highly correlated with their retention or attrition. How to model the relations of donation recurrence and donor retention and further synchronously predict these two behavioral events with a joint model are quite open problems. Finally, the presence of a large amount of censored data [9], [10], i.e., the exact attrition outcomes of some donors are unobservable or they do not perform any behaviors (donation or attrition) during our monitoring periods, imposes significant challenges in relation to this problem. Because many donors may be still in the platform in our data and most lost donors do not explicitly close their accounts when leaving, the censoring phenomenon is an inescapable concern.

To address the aforementioned issues, in this paper, we present a focused study on holistically analyzing and predicting two specific behavioral events, i.e., the *donation recurrence* and *donor retention*, in crowdfunding. That is, we aim to learn whether one donor will make donations at each time slice in the future and whether she will remain on crowdfunding platform until a future time. Specifically, we propose a Joint Deep Survival model, i.e., JDS, to jointly model these two types of behavioral events. By leveraging a deep learning framework, JDS is flexible and could integrate heterogeneous features. Also, JDS formalizes the predictions for both donation recurrence and donor retention as two collaborative goals with specific two-level predicting outputs. In addition, for modeling the censoring phenomenon and the dependence relations of different types of behaviors, we innovatively design multiple ranking constraints and incorporate them into the objective functions. Further, an alternate optimization algorithm is also proposed for effectively training JDS at two-level objectives.

The contributions of this paper can be summarized as follows:

- **Application View**. We conduct a focused study on donor retention in crowdfunding, with two specific behavioral events in charities, i.e., donation recurrence and donor retention. To the best of our knowledge, this is one of the first few attempts on comprehensively studying the donor retention problem from data-driven views in both traditional charities and crowdfunding.
- **Problem View**. We formalize the donor retention prediction as a survival analysis problem. Further, by jointly optimizing the prediction on the prior objective, i.e., donation recurrence, the complete formalization is a novel collaborative optimization problem.
- **Technical View**. We employ deep learning with ranking constraints for survival analysis which brings new insights to relevant research in this area. Also, we propose a joint deep survival model with two-level collaborative prediction outputs. Furthermore, in order to effectively train our model, we develop an alternate optimization algorithm.
- **Result View**. We collect large-scale real-world data from Kiva.org [3]. With this data, we make extensive analysis and conduct evaluation experiments whose results clearly demonstrate the effectiveness of our models towards two specific tasks.

2 Related Work

The work related to this paper is mainly studies on crowdfunding, and studies on survival analysis.

2.1 Crowdfunding

Crowdfunding is an emerging Internet-based fundraising mechanism soliciting small monetary contributions from crowd donors to help others in trouble or with dreams. Actually, more broadly speaking, crowdfunding is one specific practice of crowdsourcing [1], [11], [12], [13], [14], [15] in business or finance. In the typical crowdsourcing, researchers focus on the mechanism optimization or design, such as truth inference [16], [17], [18] and task assignments [19], [20]. Specific to crowdfunding, the topics around finance, trading or users are more concerned.

Generally speaking, the mainstream crowdfunding platforms can be classified into four categories, i.e., donation-based, reward-based, equity-based and lending-based ones [1]. Among them, the donation-based ones are becoming increasingly popular. Recently, the rapid development of crowdfunding has attracted much research attention from academics, which is mainly constructed from the *project views* [2], [21], [22] and *donor views* [4], [23], [24].

For projects, following the 'all-or-nothing' rule, the most critical concern is reaching their funding goals in time [21], [22]. Thus, some existing work focuses on predicting the project success [8], [21], [25]. For example, Lu, et al. [21] investigated the impacts of social media in crowdfunding and found the social features could help to predict the success of projects. Along this line, some researchers conducted further studies toward some advanced tasks, such as tracking the funding dynamics [2], [8], modeling the latent market states [26], recommending donors [27] and finding potential donors dynamically [22] for money-raising projects, and optimizing the settings for new-release projects [28].

From the donor viewpoint, some intelligent functions, e.g., recommending projects for donors [23], [24], [29], have been studied. For example, Zhao, et al. [23] proposed recommending project portfolios to donors with multi-objective optimization. Further, Rakesh, et al. [24] studied the group recommendation problem, i.e., recommending crowdfunding projects to a group of donors, by a proposed probabilistic generative model. In addition to the work on project recommendation, some researchers have focused on analyzing various donor behaviors, such as understanding the donation motives [7], and exploring the social communities of donors [3]. Especially, Althoff and

3. http://build.kiva.org/docs/data

Leskovec [4] explored various factors impacting donor retention in DonorsChoose.org from the statistical and empirical perspectives which was inspiring for our research. However, how to comprehensively analyze the heterogeneous factors affecting the donor retention and further predict the donor behaviors are still largely unexplored areas.

2.2 Survival Analysis

The second category of work is about survival analysis [10]. Traditionally, survival analysis is a subfield of statistics where the outcome is the time until the occurrence of an event of interest [10], such as the donor attrition in our study. Actually, customer attrition/churn [5], [6], [30], [31] is a crucial issue and has been widely studied in many commercial scenarios using traditional survival analysis, such as E-commerce, finance and services.

One of the main challenges in this context is the presence of instances whose event outcomes become unobservable after a certain time point or when some instances do not experience any event during the monitoring period, which is referred to as *censoring*. In the literature, many statistical approaches have been developed to overcome the *censoring issue* for time-to-event data, such as Cox [32], [33], [34], [35], [36] and Bayesian probabilistic algorithms [37]. Recently, many machine learning algorithms have been adopted to tackle other challenging problems that arise in the real world [10], [38], [39]. For example, Gomez-Rodriguez, et al. [40] applied survival theory to solve the network inference problems. Li, et al. [41] prospected the career paths of employees with a multi-task learning survival approach. Especially, Li, et al. [25] formulated the project success prediction in crowdfunding as a survival analysis problem and applied the censored regression approach for that. However, the problem of donor retention in crowdfunding has not been well explored.

Indeed, more advanced machine learning methods, such as ensemble learning [42], transfer learning [39], multi-task learning [41], [43] and active learning [44], have been developed to predict from censored data over the past few years [10]. Actually, these recently emerging machine learning methods as well as the traditional survival models are good at handling the structured data and modeling linear relations. However, there exists massive unstructured data and the relations of survival are quite complex in real-world scenarios. Thanks to the prevalence of deep learning, few researchers have attempted to exploit deep learning for survival analysis [31], [45], [46]. However, studies along this line rather need to be strengthened and the problem of how to utilize deep learning for survival analysis is still open. Actually, in this paper, we propose a joint deep survival model which comprises a deep learning framework with two-level prediction outputs and considers multiple ranking constraints for modeling the censored relations. The technical innovations may bring some new insights to the relevant research community. For example, our models can also be applied to survival data with modeling collaborative tasks in some other domains, such as device failure modeling in engineering, predicting student dropout [36], and prospecting the career development [41].

3 Preliminaries

In this section, we first introduce the mechanics of Kiva.org, and then provide an overview of the data collected from this website.

3.1 The Mechanics of Kiva.org

As illustrated in Fig. 1, there are mainly four types of participants in the Kiva.org community, i.e., (a) *projects*, (b) *field partners*, (c) *donors* and (d) *teams*.

The projects are in the form of money-raising campaigns/loans (borrowers only need to repay the principal without any interest) posted by field partners for the needy or suffering. A project page contains the story and the information for funding use and repayment, the details about the funding and the clients (category, goal, duration, country, poverty, etc.). Some brief information about the field partner is also included in the project page.

The field partners are local non-profit organizations, such as microfinance institutions, schools, social enterprises and charities, who are the ground links to borrowers/clients (often in developing countries) [1]. They perform their jobs mainly in two aspects. First, they need to review the funding applications from the clients and post their requests on Kiva.org. Second, they help to promote funding from donors and are responsible for repayment collection from clients. A field partner's page contains her detailed profile, such as credit, the projects she has posted in the past, etc. For both field partners and their clients, the most important issue is soliciting enough contributions from donors.

Donors are one of the most important participants in charities. Especially, since Kiva.org is a donation-based crowdfunding site, i.e., no donors receive any interest or monetary profits just their principal from the needy they help, donors essentially want to help others and focus most on the stories. Unfortunately, the donor attrition is very serious and many donors donate only once or very few times, which is exactly the issue we are focused on in this paper. A donor's page contains her donation motive, and information about projects to which she has previously donated.

For promoting donation activities, Kiva.org encourages donors to join social communities, i.e., teams. Actually, a team is made up of donors who may have similar motives, or locations [1]. One advantage of the team is that donors can collaborative in locating and funding projects. In other words, the social contacts may have a strong effect on a donor's behaviors, e.g., attrition or retention.

3.2 Data Description

Specifically, we collect the data that records the information on Kiva.org from March 2005 to July 2017. Consistent with the main participants in Kiva.org, we collect the following entities with rich information in the form of unstructured data, such as text, and structure data such as numerical, categorical, geo-spatial data, etc. Since various factors from these entities may impact the donor behaviors [4], we obtain 19 kinds of features in total.

- Projects. Kiva.org data contains 1,252,481 projects. From the project entities, we mainly extract seven kinds of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

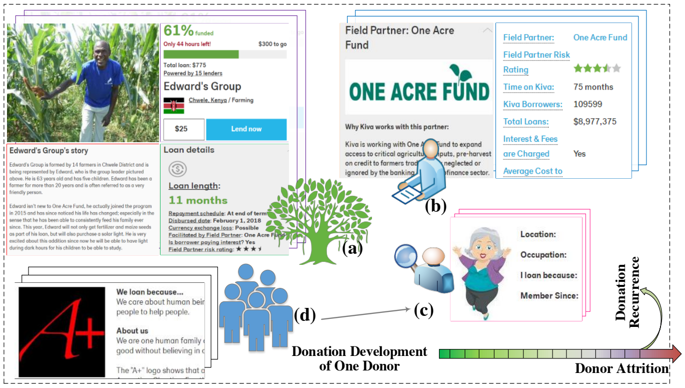


Fig. 1. An illustration of Kiva.org community.

features, they are the *funding use* in the form of text, *category*, *location*, *funding goal*, *number of funders*, *funding period*, and *repaying period*.

- Field Partners. There are 472 field partners in Kiva.org data. From these entities, four kinds of features are extracted, i.e., *credit rating*, *total number of posted projects*, *total amount of received funding for her posted projects*, and *average funding amount of each funder to her posted projects*.
- Donors. Kiva.org data contains 2,262,675 donors. From the donor entities, we mainly extract two kinds of features, i.e., *motivation* ("I donate because...") in the form of text, and *historical donations*.
- Teams. Kiva.org data contains 35,217 donor teams, from which five kinds of features are extracted, they are, *team mission* ("we donate because...") in the form of text, *category*, *number of members*, *historical funded amount*, and *number of historical projects*.
- Donations. There are 27,082,901 donation records in Kiva.org data. Considering the recent donations of one donor may impact her following behaviors [2], for each donor at a certain time interval, we construct her donation records at each timeslice in the form of a vector as one special kind of feature.

These features are heterogeneous in form. For consistency, we carefully preprocess them. Specifically, for the categorical data, we adopt the *one-hot encoding* [2], [23]. For the text data, we preliminarily take *word segment* using the National Language Toolkit (*nltk*) tool [4], and *word embedding* by representing each separate word with a 50-dimensional vector using *word2vec* [5] [47]. Additionally, all the numerical data is normalized by *Z-score* transformation [48].

4 Methodology

In this section, we first formally define the studied problem. Then, we will detail the joint deep survival model, i.e., JDS, including the framework, and optimization strategy. For better illustration, Table 1 lists mathematical notations used in this paper.

4.1 Problem Formalization

As illustrated in Fig. 2, donor i has two types of behavioral sequences, i.e., the *donation sequence* (denoted as $Y^i = \{Y_1^i, ..., Y_{T+T'}^i\}$, the black timeline) records her donations at each time slice, and the *survival/retention sequence* (denoted as $S^i = \{S_1^i, ..., S_{T+T'}^i\}$, the blue timeline) labels whether she is still staying or active on Kiva.org until this time slice. In particular, $Y_t^i = 1$ means donor i makes donations at time t (the solid circle), otherwise, she does not make any donations at that time; $S_t^i = 1$ (the blue circle) means donor i will remain on Kiva.org until time t, otherwise, $S_t^i = 0$ (the red circle) means we clearly observe that she has left this platform at or before this time. In addition, the shaded circles represent the censoring variables, i.e., the retention of this donor is not clearly known because we have not observed the occurrence of donor attrition; meanwhile, we have not observed her other behaviors such as donations at or after this time, either. Please note that we only consider the situation that donors stay on Kiva.org during the observation intervals T. The donors who have left Kiva.org in the observation intervals are not our concern. Actually, the donor retention is a *time-to-event variable* [10], [49] that often measures the length of time from some initial time until something of interest occurs, such as death in healthcare [49], and donor attrition in our study. By jointly analyzing the donation sequence and retention sequence, we have the following observations.

TABLE 1
Mathematical notations.

Notation	Description
$Y^i = \{Y_1^i, ..., Y_{T+T'}^i\}$	the donation sequence of donor i
$S^i = \{S_1^i, ..., S_{T+T'}^i\}$	the retention sequence of i
T (T')	the observation (prediction) intervals
t^*	the censoring time (censored cases)
$\widehat{Y}_t^i, t \in \{1, ..., T'\}$	the prediction of donation for i at t
$\widehat{S}_t^i, t \in \{1, ..., T'\}$	the prediction of retention for i at t
$T^i = (T_r^i, T_w^i, T_u^i, T_s^i)$	the text features corresponding to i
$X^i = (X_r^i, X_w^i, X_s^i, X_p^i)$	the other features corresponding to i
$\boldsymbol{\Theta}$	all the parameters in JDS
$\lambda_1, \lambda_2, \lambda_3, \lambda_4$	the hyper parameters

Observation 1. In our scenario, the donor attrition is permanent without turnover [6], so that $\forall t \in \{2, ..., T\} \vee t \in \{2, ..., T'\}$, if $S_t^i = 1$, then $S_{t-1}^i = 1$; $\forall t \in \{1, ..., T'-1\}$, if $S_t^i = 0$, then $S_{t+1}^i = 0$.

Observation 2. For the uncensored cases, $\max\{t\} < \min\{t'\}, (t \in \{1, ..., T\} \vee t \in \{1, ..., T'\}) \wedge Y_t^i = 1, t' \in \{1, ..., T'\} \wedge S_{t'}^i = 0$; for the censored cases, t^* is the censoring time, then $t^* - 1 = \max\{t\}, (t \in \{1, ..., T\} \vee t \in \{1, ..., T'\}) \wedge Y_t^i = 1$.

In particular, we attempt to analyze the donor behaviors with special focus on two behavioral events, i.e., whether one donor will make donations and whether she will remain on Kiva.org in the future. Thus, the studied problem can be defined.

Problem Formalization 1. For donor i, given her donation sequence $Y_t^i, t \in \{1, ..., T\}$ in the observation intervals T, and also other features including the donor features, her recent donations, i.e., T^i and X^i, our goal is to

4. http://www.nltk.org
5. https://nlp.stanford.edu/projects/glove/

6. Actually, in our setting, donors who have long donation careers can be cut into multiple instances, so that, no matter the attrition is permanent or not, the setting is reasonable because the turnover donors are treated as new ones.

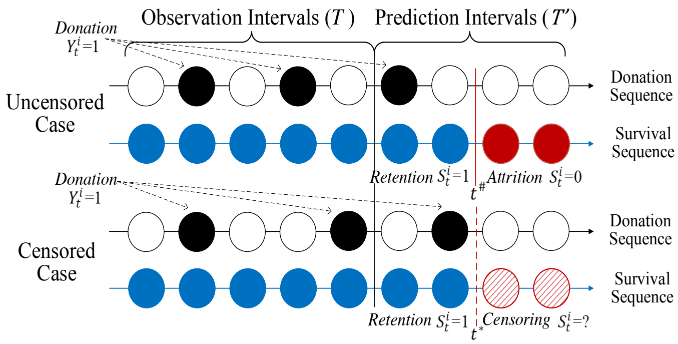


Fig. 2. Problem formalization.

analyze and predict her behaviors in the following prediction intervals T', with two special collaborative or correlated tasks:

Task1: Donation Recurrence - predicting whether donor i will donate at each time slice in T', i.e., $\widehat{Y}_t^i, t \in \{1, ..., T'\}$;

Task2: Donor Retention - predicting whether donor i will remain on Kiva.org up to each time slice in T', i.e., $\widehat{S}_t^i, t \in \{1, ..., T'\}$.

4.2 JDS Framework

Actually, donor retention can be tackled as a survival analysis problem [10], [32]. However, traditional survival models are good at handling linear relations of variables. For modeling the complex relations in donor retention and also utilizing the heterogeneous features, we propose a joint deep survival model, i.e., JDS, to jointly learn the two collaborative tasks. As illustrated in Fig. 3, JDS mainly contains three components, i.e., *Input*, *Representation* and *Prediction*. Specifically, the Input Component extracts all the heterogeneous features preliminarily, the Representation Component is used to learn the vectorial representation for each feature, and the Prediction Component gives the respective estimations for two tasks. The details of these features have been clarified in Section 3.2.

4.2.1 *Input Component*

Input Component is the fundamental part of JDS which categorizes all the heterogeneous features. Actually, there are mainly three types of features from the five entities correlated to the target donor's donation behaviors. The first type is texts, i.e., the motive of donor i in the form of "I donate because..." (denoted as T_r^i), the mission of the donor's team in the form of "we donate because..." (denoted as T_w^i), the funding use and the story of the project to which she has recently donated (denoted as T_u^i and T_s^i). We denote these text features as $T^i = (T_r^i, T_w^i, T_u^i, T_s^i)$. The second kind of important feature is the donation sequence of donor i in our observation intervals T, i.e., $Y_t^i, t \in \{1, ..., T\}$. Additionally, there are some other features including categorical ones, numerical ones and so on which we have illustrated in Section 3.2. We denote them as $X^i = (X_r^i, X_w^i, X_s^i, X_p^i)$, where each element represents the features from donor, team, project and field partner respectively.

4.2.2 *Representation Component*

Representation Component is used to learn the vectorial representation for each feature, which is one of the crucial steps. Specifically, in this step, for the donation sequence Y_t^i and other structured features X^i including the numerical, categorical, and geo-spatial ones, we take the encoding and normalization operations as referred to in Section 3.2; for the text features T^i, we employ the *Convolutional Neural Network* (CNN) to learn the vectorial representations.

Actually, the CNN [50], [51] alternates several layers of *convolution* and *p-max pooling* where each sentence is gradually summarized to a final fixed length vectorial representation. Before that, each word in the sentences is represented by a d_0-dimension vector by a pre-trained word embedding. For texts with more than one sentence, especially the project story (T_s^i), we concatenate them as one long sentence. That is $T_r^i \in \mathbb{R}^{l_r \times d_0}, T_w^i \in \mathbb{R}^{l_w \times d_0}, T_u^i \in \mathbb{R}^{l_u \times d_0}, T_s^i \in \mathbb{R}^{l_s \times d_0}$, where l_r, l_w, l_u, and l_s are the sentence lengths, i.e., word numbers of the corresponding texts. Considering the sentence lengths of $T_r^i, T_w^i, T_u^i, T_s^i$ vary a lot, we take personalized CNN structures for them. Specifically, the structures for T_r^i and T_u^i are two layers and those for T_w^i and T_s^i are three layers. Here, we introduce the first convolution-pooling operation for the sentences represented by $T_x^i \in \mathbb{R}^{l \times d_0}$, and the following deep ones for all the text features T^i are defined in a similar way.

Given the sentence matrix input $T_x^i \in \mathbb{R}^{l \times d_0}$, the narrow convolution operates on a sliding window of each k words with a kernel $k \times d_0$. In the granularity of words, $T_x^i = \{w_1, ..., w_l\}$, where w_j is the j-th word embedding vector in T_x^i. The convolution is set to obtain a new hidden sequence, i.e., $h^c = \{\overrightarrow{h}_1^c, ..., \overrightarrow{h}_{l-k+1}^c\}$, where

$$\overrightarrow{h}_j^c = \sigma(\boldsymbol{G} \cdot [w_j \oplus \cdots \oplus w_{j+k-1}] + \boldsymbol{b}), \quad (1)$$

$\boldsymbol{G} \in \mathbb{R}^{d \times kd_0}$, $\boldsymbol{b} \in \mathbb{R}^d$ are the convolution parameters, and d is the number of kernels, $\oplus$ is the operation of concatenating k word vectors into a long one, $\sigma(x)$ is a nonlinear activation function, i.e., *LeakyReLU(x)*=$\max(0, x) + negative\_shop \times \min(0, x)$, where $negative\_shop$ is a non-zero decimal.

With the convolution, each sequential k word vectors are represented by a local semantics. Then, the p-max pooling could merge the features from the convolution sequence h^c into a new global hidden sequence, i.e., $h^{cp} = \{\overrightarrow{h}_1^{cp}, ..., \overrightarrow{h}_{\lfloor (l+k-1)/p \rfloor}^{cp}\}$, where

$$\overrightarrow{h}_j^{cp} = \left[\max \begin{bmatrix} h_{pj-p+1,1}^c \\ ... \\ h_{pj,1}^c \end{bmatrix}, ..., \max \begin{bmatrix} h_{pj-p+1,d}^c \\ ... \\ h_{pj,d}^c \end{bmatrix} \right]. \quad (2)$$

Repeating the above process, more layers of convolutions and poolings are conducted to gradually summarize the global interactions of words in a sentence and finally reach the vectorial representations. Thus, for text features T^i, they are all transformed into vectorial representations.

In summary, by going through CNN for text features T^i, encoding and normalization for other structured features X^i, all the features are represented by vectors, i.e., $T_r^i \in \mathbb{R}^{d_r}$, $T_w^i \in \mathbb{R}^{d_w}$, $T_u^i \in \mathbb{R}^{d_u}$, $T_s^i \in \mathbb{R}^{d_s}$, $Y_t^i \in \mathbb{R}^T$, $X^i \in \mathbb{R}^{d_x}$, where d_x is the vectorial dimension for X^i. Then, JDS could exploit them for predictions in the next component.

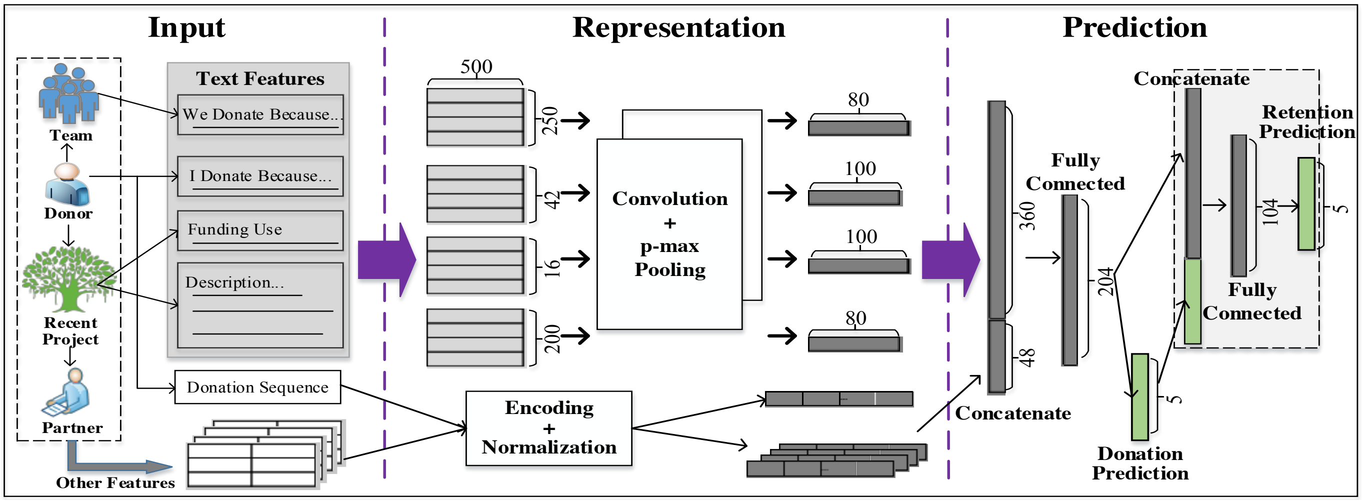


Fig. 3. The framework of JDS.

4.2.3 Prediction Component

As illustrated in Fig. 3, JDS has two-level predictions respectively aiming at the two specific tasks, i.e., the donation recurrence prediction and donor retention prediction. Please note that we argue that the prior donation behaviors, i.e., donation recurrence, do reflect one donor's decision to stay or to leave. Also, we treat Task 2 of predicting donor retention as our senior objective in this study. Thus, we design the two-level predictions, i.e., the first level is the donation recurrence prediction, and the second level is the donor retention prediction. That is, the two-level structures could exploit the predictions on donation recurrence to enhance the predictions on donor retention. Specifically, in accordance with the flows of JDS, we detail the mechanisms of predictions.

For the prediction on Task 1, with all features represented (T_r^i, T_w^i, T_u^i, T_s^i, Y_t^i, X^i), we first aggregate them by concatenation, then employ a fully-connected network [52] to learn the entire representation h_1 of all features and finally obtain the predictions on donation recurrence $\widehat{\boldsymbol{Y}}^i$ for the target donor i. That is,

$$h_1 = LeakyReLU(\boldsymbol{W_1}[T_r^i \oplus T_w^i \oplus T_u^i \oplus T_s^i \oplus Y_t^i \oplus X^i] + \boldsymbol{b_1}),$$
$$\widehat{\boldsymbol{Y}}^i = Sigmoid(\boldsymbol{W_1'} \cdot h_1 + \boldsymbol{b_1'}), \tag{3}$$

where $\widehat{\boldsymbol{Y}}^i \in \mathbb{R}^{T'}$, T' is the prediction intervals, $\boldsymbol{W_1}$, $\boldsymbol{b_1}$, $\boldsymbol{W_1'}$, $\boldsymbol{b_1'}$ are some of parameters to turn JDS. Actually, $\widehat{\boldsymbol{Y}}^i$ is a multi-prediction where each element $\widehat{Y}_t^i, t \in \{1, ..., T'\}$ is the prediction probability for donation occurring at the t-th time interval.

Thus, from the first-level prediction, we obtain the results on Task 1, i.e., whether donor i will donate at each time interval in T'. As we illustrated in the above, a donor's donation behaviors are highly correlated to her decision to stay or to leave Kiva.org. Therefore, we propose to exploit the predictions on donation recurrence to enhance the predicting performances on donor retention, i.e., involving $\widehat{\boldsymbol{Y}}^i$ to the second-level prediction. Similarly, JDS employs another concatenation and full-connected network to combine the prior representation h_1 and the first-level prediction vector $\widehat{\boldsymbol{Y}}^i$. Finally, the results on Task 2, retention prediction $\widehat{\boldsymbol{S}}^i$ for the target donor i in time intervals T' are obtained. That is,

$$h_2 = LeakyReLU(\boldsymbol{W_2} \cdot [h_1 \oplus \widehat{\boldsymbol{Y}}^i] + \boldsymbol{b_2}),$$
$$\widehat{\boldsymbol{S}}^i = Sigmoid(\boldsymbol{W_2'} \cdot h_2 + \boldsymbol{b_2'}), \tag{4}$$

where $\widehat{\boldsymbol{S}}^i \in \mathbb{R}^{T'}$, $\boldsymbol{W_2}$, $\boldsymbol{b_2}$, $\boldsymbol{W_2'}$, $\boldsymbol{b_2'}$ are some of parameters to turn JDS. The same to $\widehat{\boldsymbol{Y}}^i$, $\widehat{\boldsymbol{S}}^i$ is also a multi-prediction where each element $\widehat{S}_t^i, t \in \{1, ..., T'\}$ is the prediction probability of whether donor i will remain on Kiva.org at the t-th time interval.

In accordance with the flows of JDS, we have introduced each component. In the next subsection, we will propose an optimization strategy for effectively training JDS.

4.3 Optimization Strategy

For effectively training JDS with collaborative objectives on two tasks, we propose a holistic optimization strategy including objective functions with ranking constraints and an optimization algorithm.

4.3.1 Objective Functions with Ranking Constraints

We respectively introduce the objectives for two tasks. Firstly, for the task of predicting donation, all the donation behaviors are observable; therefore, we can formulate the objective function for Task 1 by minimizing the difference between the prediction and real record of each donor i at each time interval t, which is,

$$YL(\boldsymbol{\Theta}) = \min_{\boldsymbol{\Theta}} \sum_{i=1}^{m} ||\widehat{\boldsymbol{Y}}^i - \boldsymbol{Y}^i||^2, \tag{5}$$

where m is number of donors in training, $\boldsymbol{\Theta}$ represents all the parameters in JDS, and $||.||$ is the norm of a vector.

Differently, considering the censoring phenomenon, the donor retention is observable only before the censoring time t^*. Thus, the objective function for Task 2 can be defined as

$$SL(\boldsymbol{\Theta}) = \min_{\boldsymbol{\Theta}} \sum_{i=1}^{m} ||\widehat{\boldsymbol{S}}^i(1:t^*) - \boldsymbol{S}^i(1:t^*)||^2, \qquad (6)$$

especially for the uncensored cases, i.e., $t^* > T'$, $SL(\boldsymbol{\Theta})$ has the sample form which is similar with $YL(\boldsymbol{\Theta})$. Please note that $\boldsymbol{\Theta}$ in $SL(\boldsymbol{\Theta})$ has more parameters, i.e., $\boldsymbol{W_2}$, $\boldsymbol{b_2}$, $\boldsymbol{W_2'}$, $\boldsymbol{b_2'}$, than those in $YL(\Theta)$. For convenience, we do not distinguish between them.

For fully utilizing the censored cases and also the relations of two types of behaviors, we design ranking constraints and integrate them into two objective functions as regularization terms, i.e., $f(\boldsymbol{\Theta}, \widehat{\boldsymbol{Y}}, \widehat{\boldsymbol{S}})$. Specifically, based on Observations 1, 2, we have the following corollaries.

Corollary 1. **Non-Negativity**. The probabilities of both donation and donor retention occurrences at any time are nonnegative. That is, $\forall t \in \{1, ..., T'\}, \widehat{Y}_t^i \geq 0, \widehat{S}_t^i \geq 0$.

Corollary 2. **Non-Increment and Retention Drop**. For the donor retention sequence, the retention or survival probabilities in the prior intervals are not smaller than those of following intervals, i.e., $\forall t \in \{1, ..., T'-1\}, \widehat{S}_t^i \geq \widehat{S}_{t+1}^i$. Especially, for the uncensored intervals, if a donor leaves Kiva.org at time interval $t^{\#}$, the predicting probability of her retention should turn to smaller significantly. That is, $t^{\#} \in \{1, ..., t^*-1\}$, if $S_{t\#}^i = 1 \wedge S_{t\#+1}^i = 0$, then $\widehat{S}_{t\#}^i - \delta \geq \widehat{S}_{t\#+1}^i$, where $\delta \in (0, 1)$ is the retention drop.

Corollary 3. **Retention Priority**. For each interval (censored or uncensored), the retention probability of one donor should not be smaller than her donation probability, i.e., $\forall t \in \{1, ..., T'\}$, $\widehat{S}_t^i \geq \widehat{Y}_t^i$.

Considering these corollaries, we design the regularization terms $f(\boldsymbol{\Theta}, \widehat{\boldsymbol{Y}}, \widehat{\boldsymbol{S}})$ as follows,

$$\begin{aligned} f(\boldsymbol{\Theta}, \widehat{\boldsymbol{Y}}, \widehat{\boldsymbol{S}}) = & \overbrace{\lambda_1 \sum_{i=1}^{m} \sum_{t=1}^{T'} (-\widehat{Y}_t^i - \widehat{S}_t^i)}^{Non-Negativity} \\ & + \overbrace{\lambda_2 \sum_{i=1}^{m} (\sum_{t=1}^{T'-1} (\widehat{S}_{t+1}^i - \widehat{S}_t^i)}^{Non-Increment} + \overbrace{(\widehat{S}_{t\#+1}^i + \delta - \widehat{S}_{t\#}^i))}^{Retention\ Drop} \\ & + \overbrace{\lambda_3 \sum_{i=1}^{m} \sum_{t=1}^{T'} (\widehat{Y}_t^i - \widehat{S}_t^i)}^{Retention\ Priority} + \lambda_4 ||\Theta||_F^2, \end{aligned} \qquad (7)$$

where λ_1, λ_2, λ_3, and λ_4 are the hyper parameters which are used to balance the effects of different constraints, $||.||_F$ is the Frobennuis Norm which is used to avoid overfitting. By integrating the ranking constraints, the objective functions have the complete forms,

$$\begin{aligned} YL(\boldsymbol{\Theta}) &= \min_{\boldsymbol{\Theta}} \sum_{i=1}^{m} ||\widehat{\boldsymbol{Y}}^i - \boldsymbol{Y}^i||^2 + f(\boldsymbol{\Theta}, \widehat{\boldsymbol{Y}}, \widehat{\boldsymbol{S}}), \\ SL(\boldsymbol{\Theta}) &= \min_{\boldsymbol{\Theta}} \sum_{i=1}^{m} ||\widehat{\boldsymbol{S}}^i(1:t^*) - \boldsymbol{S}^i(1:t^*)||^2 + f(\boldsymbol{\Theta}, \widehat{\boldsymbol{Y}}, \widehat{\boldsymbol{S}}). \end{aligned} \qquad (8)$$

Then, we further develop an alternate optimization algorithm to effectively train JDS by exploiting the complete objective functions.

Algorithm 1: Alternate Optimization Algorithm.

Input: Input training data TR, initializing parameters $\boldsymbol{\Theta}$, hyper parameters and learning rates $\boldsymbol{\eta}_1$, $\boldsymbol{\eta}_2$
Output: Parameters $\boldsymbol{\Theta}$
while *not converge* **do**

$$\begin{cases} Randomly\ select\ batches\ of\ training\ instances, \\ \bigtriangledown\boldsymbol{\Theta} = \frac{\partial YL(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}, \\ \boldsymbol{\Theta} = \boldsymbol{\Theta} - \boldsymbol{\eta}_1 .* \bigtriangledown\boldsymbol{\Theta}; \end{cases}$$

$$\begin{cases} Randomly\ select\ batches\ of\ training\ instances, \\ \bigtriangledown\boldsymbol{\Theta} = \frac{\partial SL(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}, \\ \boldsymbol{\Theta} = \boldsymbol{\Theta} - \boldsymbol{\eta}_2 .* \bigtriangledown\boldsymbol{\Theta}; \end{cases}$$

return $\boldsymbol{\Theta}$.

4.3.2 *Alternate Optimization Algorithm*

As we illustrated, the donation recurrence of a donor is highly correlated to her retention in Kiva.org; also, the two-level predictions in JDS share the same feature inputs and representations. Thus, the optimization directions for two objectives are consistent, to some extent. Motivated by this characteristic, we develop an alternate optimization algorithm to jointly train JDS on two tasks. Specifically, Algorithm 1 shows the optimization. In each iteration, we minimize the objective functions with the selected training instances and optimize parameters on two objectives one after the other by back propagation with stochastic gradient descent (SGD) [52].

With the proposed optimization strategy, we enable JDS to synchronously predict the donors' multiple behaviors, i.e., donation recurrence and donor retention. For analyzing the donor behaviors and evaluating the prediction performances of JDS, we will construct massive experiments with our collected data in the next section.

5 EXPERIMENT

Specifically, we conduct the analysis and experiments from these aspects. (1) We first explore the characteristics of donation behaviors from Kiva.org data and report some findings in Section 5.1. (2) Then, we endeavor to evaluate the proposed models on two prediction tasks, including the experimental setup (Section 5.2) and various experimental results (Section 5.3).

5.1 Data Exploration

We first track the evolution of donors in Kiva.org in the entire lifecycle of our data. We note the first donation time of a donor as her coming to Kiva.org; on the other hand, if we have not observed any behaviors from one donor for more than a threshold time, e.g., one month, we denote the next time slice after her last donation as the point she left Kiva.org. Please note that, although donors may exhibit different patterns in terms of taking actions (threshold maybe different for different donors), we can not know the exact leaving points of individuals. According to the data statistics, more than 65% donors will permanently

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

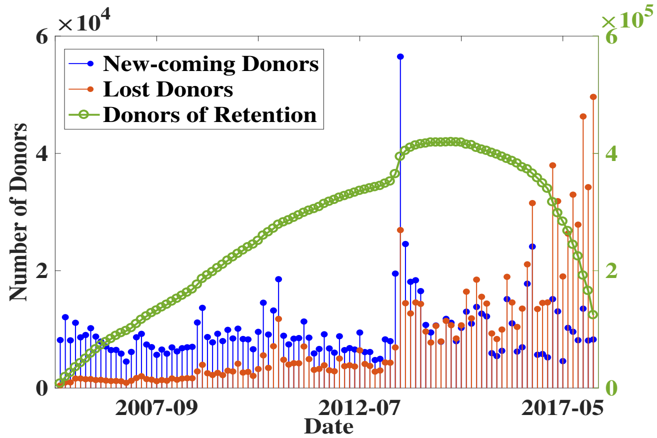


Fig. 4. The evolution of donors' retention in Kiva.org.

leave Kiva.org if we have not observed any behaviors of them in one month. In addition, even though a few donors may come back to Kiva.org more than one month later, the models may loss the abilities of capturing the sequence dependence for such a long time. For those case, models can treat the turnover donors as new ones. Thus, in this study, we empirically set the threshold time as one month. The details of the statistical results are shown in Fig. 4. Specifically, the blue stem represents the number of new-coming donors, and the red stem represents the number of lost donors at each month. We can see that, in the first several years, Kiva.org attracted more donors than lost ones. However, in recent years, more donors have left Kiva.org and the attrition is very serious. The green line represents the retention of donors in Kiva.org, which also indicates similar observations.

Then, we report the survival or retention of donors in Kiva.org in Fig. 5. First, Fig. 5(a) shows the fraction of donors with different survival days in the entire data. Actually, both the density proportion and cumulative proportion indicate the 'long-tailed effect' of donor retention [53], i.e., many donor retentions are not long. Going a step further, we respectively categorize donors based on whether they have claimed their $Motives$ in their profile pages and whether they have joined any $Teams$. The respective distributions are shown in Fig. 5(b) and Fig. 5(c). The 'long-tailed effect' can be observed in each category of donor retention. Additionally, both in Fig. 5(b) and Fig. 5(c), the retention of different donors is clearly different, which demonstrates that both the motives and social contacts of donors do affect their survival or retention.

Further, we analyze the vitality of charity activities in Kiva.org. Specifically, Fig. 6(a) shows what fraction of donors make how many donations within the entire data period. About 35% of donors make exactly one donation and never return, and only 12% of donors make more than 25 donations in Kiva.org. This preliminary exploration tells us that the donor attrition is high and a quite large fraction of donors is lost after several donations in Kiva.org. This is where we focus our attention in this paper, i.e., analyzing the influence factors of and further predicting the donation recurrence and donor retention.

The comparative statistics of donors with(out) declared motives or teams are shown in Fig. 6(b) and 6(c). Specifically, from Fig. 6(b), we find that distributions of the respective fractions are quite different, i.e., donors with declared motives are more likely to have longer donation careers and make more donations. Similar results are also observed from the comparison between donors who belong to teams and those who do not. We conclude that both the donors' profiles, such as motives, and social contacts, such as teams, do have great effects on their donations. The findings are consistent with the reported results in [4]. Limited by space, we do not report more results on other factors, such as the recently donated projects, or field partners. In fact, more factors and effects have been clearly analyzed in [4].

In Fig. 7, we plot the number of donor donations with respect to their lifecycles, i.e., retentions. A positive correlation (*Pearson correlation coefficient*, 0.681) can be observed which implies the rationality of joint learning on two prediction tasks.

5.2 Experimental Setup

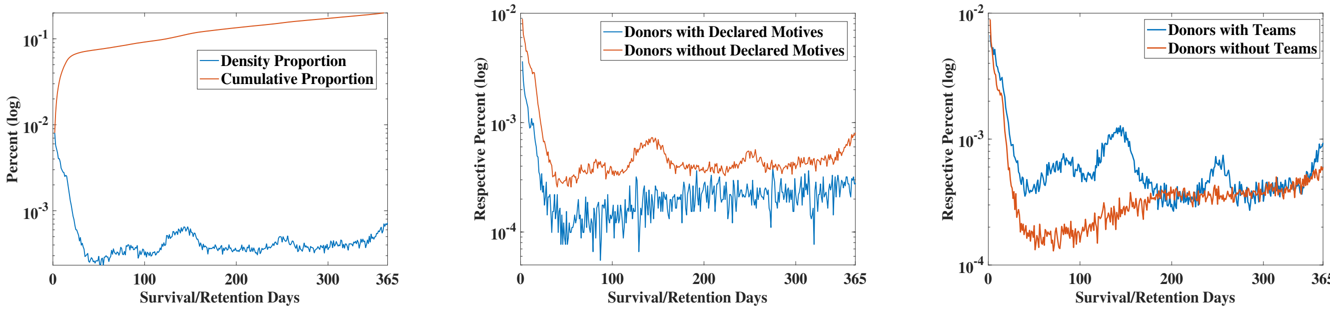
To evaluate the prediction performances of JDS on two tasks, we conduct experiments with the Kiva.org data.

5.2.1 Data Partitioning

For comparable evaluations, we remove the donors who have fewer than two donations. After that, we still have 383,146 projects, 386 field partners, 42,768 donors, 26,958 teams and 1,193,148 donation records. In practice, we set the time slice as each week. At each time slice, the variables are constructed by merging those corresponding features of each day belonging to this slice. We have two settings for the data partition. In the first way, the observation intervals are four slices (28 days), i.e., $T = 4$, and the prediction intervals are five slices, i.e., $T' = 5$, which is denoting as '4-5'partitioning. In the second way, $T = 6$ and $T' = 3$, which is denoting as '6-3'partitioning. Thus, these donors who have long donation careers may be cut into multiple instances. In summary, we have 501,778 instances in total. We randomly select 80% of them as training instances, and the remaining 20% of instances are used to test.

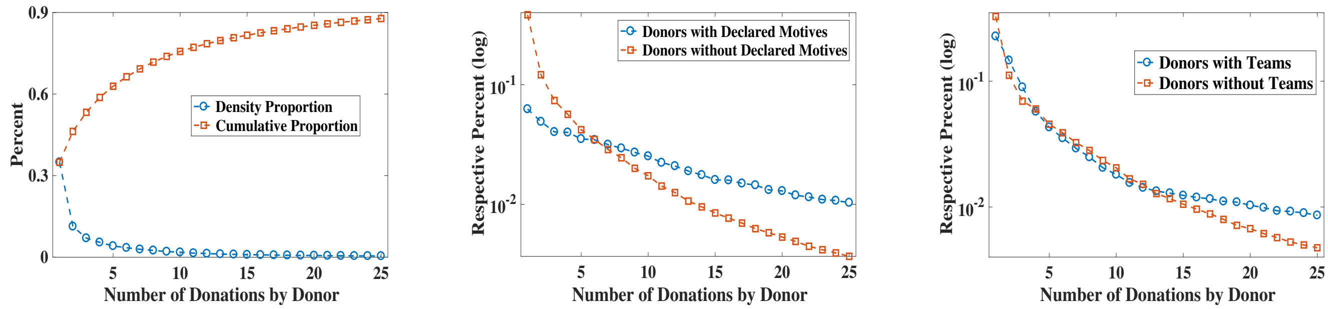
5.2.2 Parameter Setting

The dimensions of all the feature representations at each layer in JDS are labeled in Fig. 3. In Representation Component, the kernel sizes are set as [3,50]*100 and [3,100]*50 for the first and second layer convolutions in processing the donor texts. Specifically, [3,50] is the kernel matrix size and 100 is the number of kernels. In the same way, the kernels are respectively set as ([6,50]*100, [5,100]*50, [2,50]*40), ([2,50]*100, [2,100]*50) and ([6,50]*100, [5,100]*50, [2,50]*40) for team texts, funding use, and descriptions. Also, the pooling windows p are correspondingly set as ([4,1], [4,1]), ([5,1], [5,1], [4,1]), ([3,1], [2,1]), ([5,1], [5,1], [3,1]). The $negative\_slope$ in the activation function is set as 0.2. We follow [54] and randomly initialize all matrix parameters at each layer with a uniform distribution in the range of $(-\sqrt{6/(nin+nout)}, \sqrt{6/(nin+nout)})$, where nin and $nout$ are the numbers of input and output features at each layer in JDS. When training JDS, we summarize the instances from one donor into a mini batch. The dropout



(a) Survival days of all donors.
(b) Survival days of donors with(out) motives.
(c) Donations of donors in teams or not.

Fig. 5. Survival/Retention analysis.



(a) Donations of all donors.
(b) Donations of donors with(out) motives.
(c) Donations of donors in teams or not.

Fig. 6. Donation behavior analysis.

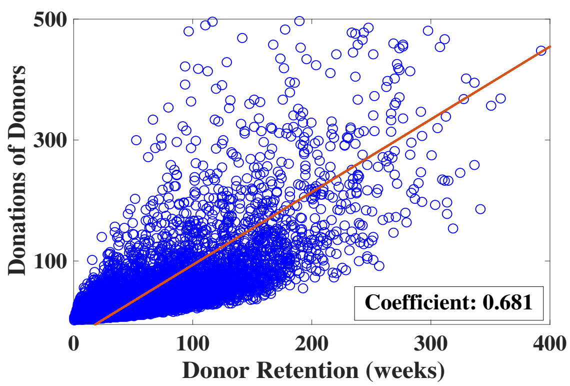


Fig. 7. Correlation (Donation and Retention).

probability is 0.8, and all the parameters are tuned when training. Without special illustration, the hyper parameters λ_1, λ_2, λ_3, and λ_4 are respectively set as 0.07, 0.07, 0.02, 10^{-5} and the retention drop δ is set as 0.1 for their best performances in the empirical validations.

5.2.3 Evaluation Metrics

Due to the presence of censored data, we adopt a widely-used evaluation metric, i.e., the concordance index (*C-index*), in survival analysis [10], [41], [43] to evaluate the prediction performances. Actually, the *C-index* computes the consistency between prediction and reality. Specifically, we respectively define metrics for two prediction tasks, i.e., C_y for donation recurrence and C_s for donor retention at each time interval t.

$$
\begin{aligned}
C_y &= \frac{1}{M_1}\sum_{(i=1\wedge Y_t^i=0)}^{m}\ \sum_{(k=1\wedge Y_t^k=1)}^{m} I[\widehat{Y}_t^k > \widehat{Y}_t^i], t\in\{1,...,T'\},\\
C_s &= \frac{1}{M_2}\sum_{(i=1\wedge S_t^i=0)}^{m}\ \sum_{(k=1\wedge S_t^k=1)}^{m} I[\widehat{S}_t^k > \widehat{S}_t^i], t\in\{1,...,t^*\},
\end{aligned}
\tag{9}
$$

where M_1 and M_2 are the numbers of comparable pairs in computing, $I[x]$ is an indicator function that equals to 1 if x is true and equals to 0 otherwise. For both the two tasks, JDS predicts the occurrence probabilities (0,1) of two types of behavioral events; thus, we can also evaluate the tasks from classification view with best thresholds cutting the probabilities. Specifically, we borrow the widely-used metric $F1$ measure [55] from classification evaluations.

5.2.4 Comparison Methods

Our complete method which jointly learns the two tasks and predicts them simultaneously is denoted as JDS in the experiments. Additionally, considering the difference between two tasks, we respectively select representative benchmark methods for them.

Baselines in Task 1. For predicting donation recurrence, we aim at predicting the probabilities for each time slice in intervals T'. Consequently, multi-task learning approaches are very suitable:

- **M-L2,1** [41], [56], which is a standard multi-task learning model with $\ell_{2,1}$ norm penalty.
- **M-LASSO** [41], [56], which is a standard multi-task learning model with LASSO norm penalty.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

TABLE 2
Performance of predicting donation recurrence on two metrics (%). left: C_y, right: $F1$. '4-5' partitioning data.

Methods.\ T'	1	2	3	4	5	[1-5]	1	2	3	4	5	[1-5]
M-L2,1	52.154	53.849	80.716	79.951	80.866	59.646	31.822	28.627	17.377	16.696	16.866	22.278
M-LASSO	52.996	54.619	85.395	84.876	85.541	61.408	31.853	28.668	19.998	20.981	20.709	24.442
JDS-Y	72.315	71.662	**83.565**	**84.443**	**82.108**	74.378	**57.131**	**53.111**	**34.576**	37.273	32.280	42.874
JDS	**73.593**	**73.383**	83.263	83.419	81.522	**75.423**	56.834	52.888	34.327	**38.255**	**32.93**	**43.047**

TABLE 3
Performance of predicting donor retention on two metrics (%). left: C_s, right: $F1$. '4-5' partitioning data.

Methods.\ T'	1	2	3	4	5	[1-5]	1	2	3	4	5	[1-5]
M-L2,1	64.824	64.655	88.770	87.670	86.111	76.227	91.160	88.890	57.723	52.400	45.106	67.056
M-LASSO	65.093	64.499	88.933	87.618	86.626	76.353	91.155	88.897	58.687	53.141	47.110	67.798
COX	–	–	–	–	–	42.913	–	–	–	–	–	19.361
Logistic	–	–	–	–	–	53.433	–	–	–	–	–	44.648
Log-Logistic	–	–	–	–	–	70.016	–	–	–	–	–	49.205
Tobit	–	–	–	–	–	52.016	–	–	–	–	–	45.048
CDT	69.847	68.875	86.884	85.550	84.183	77.515	91.145	88.880	65.597	59.037	48.622	70.656
JDS-S	78.357	**73.100**	86.828	85.290	83.863	80.503	92.149	89.936	64.195	57.437	49.717	70.687
JDS	**78.662**	73.091	**91.699**	**90.818**	**89.180**	**83.201**	**92.161**	**90.055**	**67.918**	**62.343**	**54.301**	**73.356**

TABLE 4
Performance of predicting donation recurrence on two metrics (%). left: C_y, right: $F1$. '6-3' partitioning data.

Methods.\ T'	1	2	3	[1-3]	1	2	3	[1-3]
M-L2,1	77.998	76.950	77.939	77.608	20.142	16.715	19.122	18.660
M-LASSO	81.134	76.805	77.628	78.471	21.477	17.541	20.107	19.708
JDS-Y	**83.448**	84.013	**82.497**	**83.332**	**42.576**	45.846	37.428	41.950
JDS	82.256	**84.622**	82.367	83.221	42.288	**47.489**	**37.922**	**42.566**

TABLE 5
Performance of predicting donor retention on two metrics (%). left: C_s, right: $F1$. '6-3' partitioning data.

Methods.\ T'	1	2	3	[1-3]	1	2	3	[1-3]
M-L2,1	81.624	82.024	80.990	81.572	56.442	52.540	47.252	52.780
M-LASSO	82.471	81.412	78.618	80.984	56.469	52.592	46.809	51.957
COX	–	–	–	51.541	–	–	–	19.002
Logistic	–	–	–	57.108	–	–	–	42.289
Log-Logistic	–	–	–	54.254	–	–	–	46.700
Tobit	–	–	–	60.897	–	–	–	57.834
CDT	85.376	82.867	83.044	83.851	65.768	60.184	46.460	57.470
JDS-S	85.123	83.790	83.601	84.228	**66.168**	**61.058**	48.879	58.702
JDS	**89.751**	**88.147**	**87.256**	**88.703**	66.157	60.177	**52.468**	**59.601**

- **JDS-Y** is a variant of JDS, which only has a one-level prediction on donation recurrence by removing the follow-on prediction structure on donor retention.

Baselines in Task 2. Both M-L2,1 and M-LASSO can also be directly applied to Task 2 by only modeling the uncensored data. Further, for exploiting both the uncensored and censored data, we borrow some representative models from survival analysis:

- **COX** [32], which is one of the most widely-used semi-parametric survival models. COX models the hazard function in exp proportional fashion and relates to a baseline hazard function.
- **Logistic** [25], **Log-Logistic** [25], which are parametric survival models with logistic or Log-logistic distributions respectively.
- **Tobit** [10], [57], also called a censored regression model, is designed to estimate linear relationships between variables when there is censoring in the dependent variables.
- **CDT** [41] is a multi-task learning framework for survival analysis and could exploit Non-Increment and Retention Drop constraints.
- **JDS-S** is another variant of JDS by removing the first-level prediction structure on donation recurrence and directly connecting the final prediction to the Representation Component.

All these methods use the same features and representations and also are trained with parameters which perform their best on the training data. It is worth noting that COX, Logistic, Log-Logistic and Tobit are the traditional survival models with only one output value which is the estimated relative priority of event occurrence starting from an observable time. Thus, these models could only predict the probabilities for entire prediction intervals, i.e., T', rather than each time slice. Differently, M-L2,1, M-LASSO, JDS-Y, JDS-S and JDS are all equipped with multi-task outputs so that they could predict for each time slice.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

5.3 Experimental Results

We respectively report the experimental results from the joint training, the prediction performances on two tasks, and the study of hyper parameters on evaluating the ranking constraints.

5.3.1 Joint Training Result

We respectively train the JDS, JDS-Y, and JDS-S and record their performances on two tasks by computing C_y or C_s in each iteration. Limited by the space, we only report the results on the '4-5'partitioning data in this part. The results are shown in Fig. 8. Clearly, from the comparisons between JDS and JDS-Y or JDS-S, we can see that JDS can be optimized more quickly and then converge. Especially, the JDS model has a more significant advantage in relation to the second task. Thus, we conclude that the donation recurrence and donor retention are highly correlated and these two tasks could learn from each other when optimizing parameters. Also, the results indicate the effectiveness of JDS structure and the joint optimization strategy.

5.3.2 Prediction Performance

Specifically, the prediction performances of all comparison methods on the two tasks with '4-5'partitioning data are respectively shown in Table 2 and Table 3. Firstly, we pay attention to the performance on Task 1. Both on C_y and F_1, our model, either JDS or JDS-Y, performs significantly better than the traditional multi-task models. Specifically, compared with M-L2,1 and M-LASSO in the entire prediction intervals, JDS or JDS-Y respectively performs improvements with more than 20% and 70% on C_y and F_1 measures. These results clearly demonstrate the JDS model's abilities in integrating heterogeneous features and modeling their complex relations for better predicting the donation recurrence.

Then, we turn to the results on Task 2. In most cases, JDS or JDS-S performs best with at least 9% and 4% improvements on C_s and F_1 measures. More specifically, different from the results on Task 1, JDS almost dominates JDS-S on this task, which indicates the help from incorporating the donation recurrence prediction into Task 2 and the utility of joint optimization. The traditional models, i.e., COX, Logistic, Log-Logistic and Tobit, could not provide competitive results due to their insufficiency in handling heterogeneous features and the complex relations in our study. The standard multi-task models, i.e., M-L2,1 and M-LASSO, could also not provide satisfactory results, too, because they do not have the ability to model the censored data. Relatively speaking, CDT is competitive since it takes the advantages of both multi-task learning and modeling censored data.

One special finding in Table 2 and Table 3 is that almost all methods do not perform well on C_y at the first two prediction intervals. According to our observation and analysis, the probable reason is that the occurrence of the two behavioral events (i.e., donation recurrence and donor retention) at those intervals are very unbalance.

Further, Table 4 and Table 5 show the results with with '6-3'partitioning data, where the referred phenomenon is not observed. In general, the results in Table 4 and Table 5 are similar with those in Table 2 and Table 3, which further certify the performance of our proposed models.

5.3.3 Study of Features

For evaluating and comparing the utilities of features from different entities (i.e., donor, team, project, partner and donation) on prediction, we respectively obtain the results using JDS on two tasks when removing one specific kind entity of features at each round. Limited by the space, we only report the results on the entire prediction intervals with '4-5'partitioning data in this part. Specifically, the results are shown in Fig. 9. Preliminarily, we can see that separately removing any kind entity of features will bring varying degrees of loss. Relatively, the individual importance or effects of different kind entity of features on the two prediction tasks is: Donation>Project>Donor>Partner>Team.

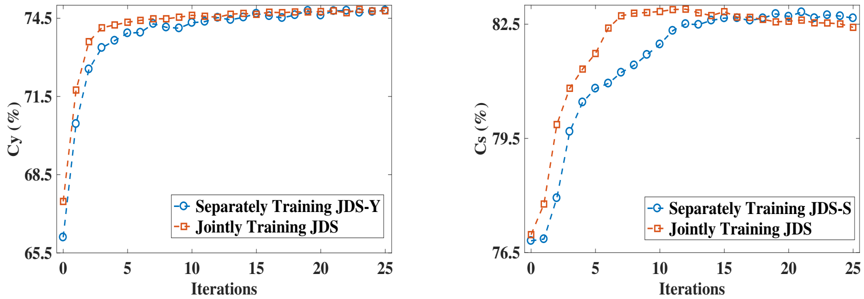
5.3.4 Study of Hyper Parameters

We also evaluate the effectiveness of ranking constraints by studying the hyper parameters on two tasks. Specifically, we report the results on two tasks with '4-5'partitioning data by changing one parameter gradually and keeping the others invariant. The results are respectively shown in Fig. 10(a), 10(b), 10(c). With the parameters increasing (λ_1 and λ_2 are from 0.05 to 0.09, λ_3 is from 0.01 to 0.05), the performance of the JDS model in relation to the two tasks increases first and then decreases. The best settings are λ_1=0.07, λ_2=0.07 and λ_3=0.02 respectively. In fact, the hyper parameters reflect the importance of the multiple ranking constraints when modeling censoring variables and the relations of behavioral events.

6 Conclusion

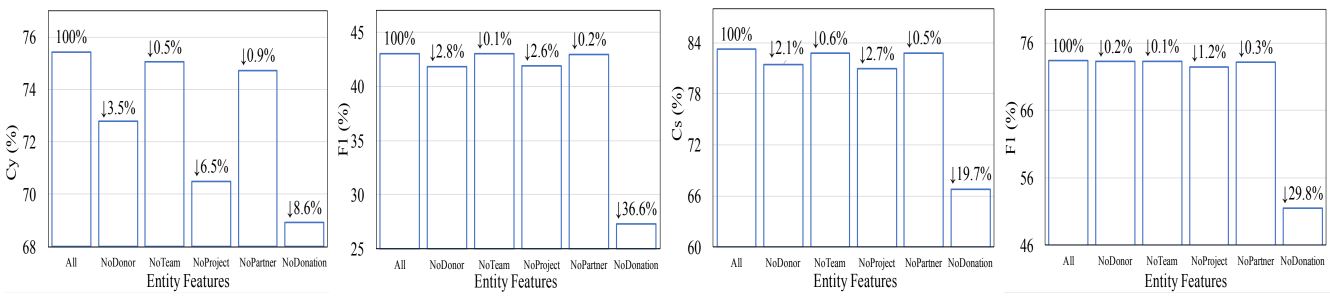
In this paper, we presented a focused study on prospecting the donation careers in crowdfunding. By collecting and analyzing large-scale real-world data, we specifically formalized two predicting tasks on donation recurrence and donor retention. Then, using a data-driven method, we proposed a Joint Deep Survival model, i.e., JDS, which could integrate heterogeneous features to jointly model the donation recurrence and donor retention. Additionally, we designed multiple innovative constraints and incorporated them into objective functions for modeling the censoring phenomenon and dependence relations of different behaviors when training JDS. In experiments, we analyzed the donations in crowdfunding and validated the prediction performances of JDS on two tasks from various aspects. The experimental results clearly demonstrated the effectiveness of our proposed models for analyzing and predicting the behavioral events, i.e., donation recurrence and donor retention.

Our study may bring some new insights from the application view of crowdfunding and the technical view of exploiting deep learning for survival analysis to the research communities. In the future, we will apply and improve our models for other scenarios, such as traditional charity activities, especially applied to survival data with modeling collaborative tasks in some other domains, such as device failure modeling in engineering, predicting student dropout, and prospecting the career development.



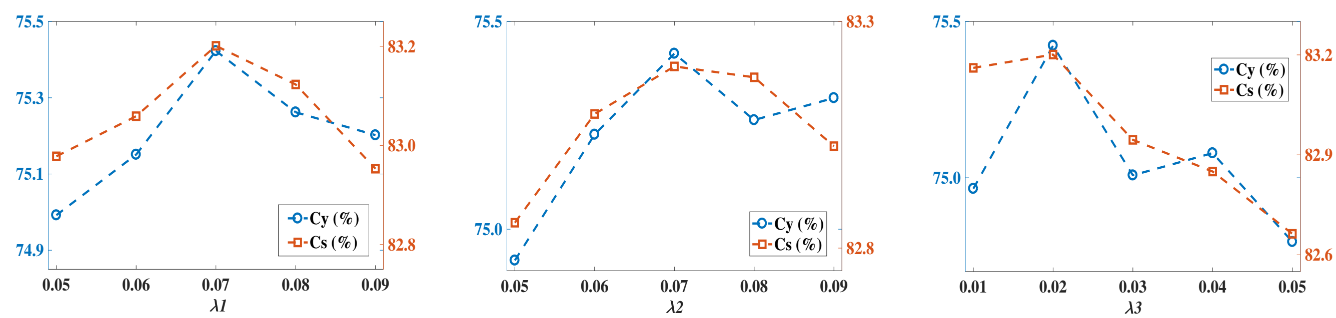
(a) Training performances on Task 1. (b) Training performances on Task 2.

Fig. 8. Training performances on two Tasks.



(a) Results of C_y on Task 1. (b) Results of $F1$ on Task 1. (c) Results of C_s on Task 2. (d) Results of $F1$ on Task 2.

Fig. 9. The performance loss with removing different entity features on two prediction tasks.



(a) Effects of λ_1 on prediction performances. (b) Effects of λ_2 on prediction performances. (c) Effects of λ_3 on prediction performances.

Fig. 10. The effects of hyper parameters on two prediction tasks.

REFERENCES

[1] H. Zhao, Y. Ge, Q. Liu, G. Wang, E. Chen, and H. Zhang, "P2p lending survey: Platforms, recent advances and prospects," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 8, no. 6, p. 72, 2017.

[2] H. Zhao, H. Zhang, Y. Ge, Q. Liu, E. Chen, H. Li, and L. Wu, "Tracking the dynamics in crowdfunding," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 625–634.

[3] J. Choo, D. Lee, B. Dilkina, H. Zha, and H. Park, "To gather together for a better world: Understanding and leveraging communities in micro-lending recommendation," in *Proceedings of the 23rd international conference on World wide web*. ACM, 2014, pp. 249–260.

[4] T. Althoff and J. Leskovec, "Donor retention in online crowdfunding communities: A case study of donorschoose. org," in *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2015, pp. 34–44.

[5] L. J. Rosenberg and J. A. Czepiel, "A marketing approach for customer retention," *Journal of consumer marketing*, vol. 1, no. 2, pp. 45–51, 1984.

[6] H. Datta, B. Foubert, and H. J. Van Heerde, "The challenge of retaining customers acquired with free trials," *Journal of Marketing Research*, vol. 52, no. 2, pp. 217–234, 2015.

[7] Y. Liu, R. Chen, Y. Chen, Q. Mei, and S. Salib, "I loan because...: Understanding motivations for pro-social lending," in *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, 2012, pp. 503–512.

[8] Y. Zhang, Y. Xiong, X. Kong, and Y. Zhu, "Netcycle: Collective evolution inference in heterogeneous information networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1365–1374.

[9] B. Vinzamuri, Y. Li, and C. K. Reddy, "Pre-processing censored survival data using inverse covariance matrix based calibration," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 10, pp. 2111–2124, 2017.

[10] P. Wang, Y. Li, and C. K. Reddy, "Machine learning for survival analysis: A survey," *arXiv preprint arXiv:1708.04649*, 2017.

[11] H. Garcia-Molina, M. Joglekar, A. Marcus, A. Parameswaran, and V. Verroios, "Challenges in data crowdsourcing," *IEEE Transactions*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2019.2906199, IEEE Transactions on Knowledge and Data Engineering

on Knowledge and Data Engineering, vol. 28, no. 4, pp. 901–911, 2016.

[12] G. Li, Y. Zheng, J. Fan, J. Wang, and R. Cheng, "Crowdsourced data management: Overview and challenges," in *Proceedings of the 2017 ACM International Conference on Management of Data*. ACM, 2017, pp. 1711–1716.

[13] A. I. Chittilappilly, L. Chen, and S. Amer-Yahia, "A survey of general-purpose crowdsourcing techniques," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 9, pp. 2246–2266, 2016.

[14] G. Li, J. Wang, Y. Zheng, and M. J. Franklin, "Crowdsourced data management: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 9, pp. 2296–2319, 2016.

[15] G. Li, J. Wang, Y. Zheng, J. Fan, and M. J. Franklin, *Crowdsourced Data Management: Hybrid Human-Machine Data Management*. Springer, 2018.

[16] Y. Zheng, R. Cheng, S. Maniu, and L. Mo, "On optimality of jury selection in crowdsourcing," in *Proceedings of the 18th International Conference on Extending Database Technology, EDBT 2015*. OpenProceedings. org., 2015.

[17] Y. Zheng, G. Li, and R. Cheng, "Docs: a domain-aware crowdsourcing system using knowledge bases," *Proceedings of the VLDB Endowment*, vol. 10, no. 4, pp. 361–372, 2016.

[18] Y. Zheng, G. Li, Y. Li, C. Shan, and R. Cheng, "Truth inference in crowdsourcing: Is the problem solved?" *Proceedings of the VLDB Endowment*, vol. 10, no. 5, pp. 541–552, 2017.

[19] H. Hu, Y. Zheng, Z. Bao, G. Li, J. Feng, and R. Cheng, "Crowdsourced poi labelling: Location-aware result inference and task assignment," in *Data Engineering (ICDE), 2016 IEEE 32nd International Conference on*. IEEE, 2016, pp. 61–72.

[20] Y. Zheng, J. Wang, G. Li, R. Cheng, and J. Feng, "Qasca: A quality-aware task assignment system for crowdsourcing applications," in *Proceedings of the 2015 ACM SIGMOD international conference on management of data*. ACM, 2015, pp. 1031–1046.

[21] C.-T. Lu, S. Xie, X. Kong, and P. S. Yu, "Inferring the impacts of social media on crowdfunding," in *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, 2014, pp. 573–582.

[22] H. Zhang, H. Zhao, Q. Liu, T. Xu, E. Chen, and X. Huang, "Finding potential lenders in p2p lending: A hybrid random walk approach," *Information Sciences*, vol. 432, pp. 376–391, 2018.

[23] H. Zhao, Q. Liu, G. Wang, Y. Ge, and E. Chen, "Portfolio selections in p2p lending: A multi-objective perspective," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 2075–2084.

[24] V. Rakesh, W.-C. Lee, and C. K. Reddy, "Probabilistic group recommendation model for crowdfunding domains," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. ACM, 2016, pp. 257–266.

[25] Y. Li, V. Rakesh, and C. K. Reddy, "Project success prediction in crowdfunding environments," in *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. ACM, 2016, pp. 247–256.

[26] H. Zhao, Q. Liu, H. Zhu, Y. Ge, E. Chen, Y. Zhu, and J. Du, "A sequential approach to market state modeling and analysis in online p2p lending," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 1, pp. 21–33, 2018.

[27] J. An, D. Quercia, and J. Crowcroft, "Recommending investors for crowdfunding projects," in *Proceedings of the 23rd international conference on World wide web*. ACM, 2014, pp. 261–270.

[28] Q. Liu, G. Wang, H. Zhao, C. Liu, T. Xu, and E. Chen, "Enhancing campaign design in crowdfunding: A product supply optimization perspective," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, ser. IJCAI'17. AAAI Press, 2017, pp. 695–702.

[29] J. Choo, C. Lee, D. Lee, H. Zha, and H. Park, "Understanding and promoting micro-finance activities in kiva. org," in *Proceedings of the 7th ACM international conference on Web search and data mining*. ACM, 2014, pp. 583–592.

[30] T. Iwata, K. Saito, and T. Yamada, "Recommendation method for improving customer lifetime value," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 9, pp. 1254–1263, 2008.

[31] H. Jing and A. J. Smola, "Neural survival recommender," in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM, 2017, pp. 515–524.

[32] D. R. Cox, "Regression models and life-tables," in *Breakthroughs in statistics*. Springer, 1992, pp. 527–541.

[33] K. Kapoor, M. Sun, J. Srivastava, and T. Ye, "A hazard based approach to user return time prediction," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 1719–1728.

[34] P. Yadav, M. Steinbach, L. Pruinelli, B. Westra, C. Delaney, V. Kumar, and G. Simon, "Forensic style analysis with survival trajectories," in *Data Mining (ICDM), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1069–1074.

[35] G. J. Simon, P. J. Caraballo, T. M. Therneau, S. S. Cha, M. R. Castro, and P. W. Li, "Extending association rule summarization techniques to assess risk of diabetes mellitus," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 1, pp. 130–141, 2015.

[36] S. Ameri, M. J. Fard, R. B. Chinnam, and C. K. Reddy, "Survival analysis based framework for early prediction of student dropouts," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 2016, pp. 903–912.

[37] M. J. Fard, P. Wang, S. Chawla, and C. K. Reddy, "A bayesian perspective on early stage event prediction in longitudinal data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 12, pp. 3126–3139, 2016.

[38] L. Yu, P. Cui, F. Wang, C. Song, and S. Yang, "From micro to macro: Uncovering and predicting information cascading process with behavioral dynamics," in *Data mining (ICDM), 2015 IEEE international conference on*. IEEE, 2015, pp. 559–568.

[39] Y. Li, L. Wang, J. Wang, J. Ye, and C. K. Reddy, "Transfer learning for survival analysis via efficient l2, 1-norm regularized cox regression," in *Data Mining (ICDM), 2016 IEEE 16th International Conference on*. IEEE, 2016, pp. 231–240.

[40] M. Gomez-Rodriguez, J. Leskovec, and B. Schölkopf, "Modeling information propagation with survival theory," in *International Conference on Machine Learning*, 2013, pp. 666–674.

[41] H. Li, Y. Ge, H. Zhu, H. Xiong, and H. Zhao, "Prospecting the career development of talents: A survival analysis perspective," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 917–925.

[42] T. Hothorn, B. Lausen, A. Benner, and M. Radespiel-Tröger, "Bagging survival trees," *Statistics in medicine*, vol. 23, no. 1, pp. 77–91, 2004.

[43] Y. Li, J. Wang, J. Ye, and C. K. Reddy, "A multi-task learning formulation for survival analysis," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 1715–1724.

[44] B. Vinzamuri, Y. Li, and C. K. Reddy, "Active learning based survival regression for censored data," in *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*. ACM, 2014, pp. 241–250.

[45] M. Luck, T. Sylvain, H. Cardinal, A. Lodi, and Y. Bengio, "Deep learning for patient-specific kidney graft survival analysis," *arXiv preprint arXiv:1705.10245*, 2017.

[46] S. Yousefi, F. Amrollahi, M. Amgad, C. Dong, J. E. Lewis, C. Song, D. A. Gutman, S. H. Halani, J. E. V. Vega, D. J. Brat *et al.*, "Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models," *Scientific Reports*, vol. 7, no. 1, p. 11707, 2017.

[47] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[48] C. Cheadle, M. P. Vawter, W. J. Freed, and K. G. Becker, "Analysis of microarray data using z score transformation," *The Journal of molecular diagnostics*, vol. 5, no. 2, pp. 73–81, 2003.

[49] J. P. Klein and M. L. Moeschberger, *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media, 2005.

[50] Y. Kim, "Convolutional neural networks for sentence classification," *arXiv preprint arXiv:1408.5882*, 2014.

[51] A. Severyn and A. Moschitti, "Learning to rank short text pairs with convolutional deep neural networks," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2015, pp. 373–382.

[52] R. Hecht-Nielsen *et al.*, "Theory of the backpropagation neural network." *Neural Networks*, vol. 1, no. Supplement-1, pp. 445–448, 1988.

[53] A. B. Downey, "Evidence for long-tailed distributions in the internet," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*. ACM, 2001, pp. 229–241.

[54] G. B. Orr and K.-R. Müller, *Neural networks: tricks of the trade*. Springer, 2003.

[55] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation," 2011.

[56] J. Zhou, L. Yuan, J. Liu, and J. Ye, "A multi-task learning formulation for predicting disease progression," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2011, pp. 814–822.

[57] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica: journal of the Econometric Society*, pp. 24–36, 1958.

Hongke Zhao received the Ph.D. degree from the University of Science and Technology of China, Hefei, China. He is currently a faculty member of the College of Management and Economics, Tianjin University. His research interest includes data mining, knowledge management and Internet finance-based applications, such as Crowdfunding and P2P lending.

He has published more than 20 papers in refereed journals and conference proceedings, such as *ACM Transactions on Intelligent Systems and Technology*, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *IEEE Transactions on Big Data*, *Information Sciences*, ACM SIGKDD, IJCAI, AAAI and IEEE ICDM. He was the program committee member at ACM SIGKDD, AAAI, etc.

Binbin Jin received the B.E. degree in computer science from the University of Science and Technology of China (USTC), Hefei, China, 2016. He is currently working toward the Ph.D. degree in the Department of Computer Science and Technology, USTC. His research interests include data mining, deep learning, survival analysis, and applications in Community Question Answering (CQA) and Internet finance-based applications, such as crowdfunding.

Qi Liu received the Ph.D. degree in computer science from University of Science and Technology of China (USTC). He is an associate professor with the USTC. His general area of research is data mining and knowledge discovery. He has published prolifically in refereed journals and conference proceedings, e.g., the *IEEE Transactions on Knowledge and Data Engineering*, the *ACM Transactions on Information Systems*, the *ACM Transactions on Knowledge Discovery from Data*, the *ACM Transactions on Intelligent Systems and Technology*, ACM SIGKDD, IJCAI, AAAI, IEEE ICDM, SDM, and ACM CIKM. He has served regularly on the program committees of a number of conferences, and is a reviewer for the leading academic journals in his fields. He received the ICDM 2011 Best Research Paper Award and the Best of SDM 2015 Award. He is a member of ACM and the IEEE.

Yong Ge received the Ph.D. degree in information technology from the State University of New Jersey, Rutgers, in 2013. He is an assistant professor of management information systems with the University of Arizona. His research interests include data mining and business analytics. He received the ICDM-2011 Best Research Paper Award. He has published prolifically in refereed journals and conference proceedings, such as the *IEEE Transactions on Knowledge and Data Engineering*, the *ACM Transactions on Information Systems*, the *ACM Transactions on Knowledge Discovery from Data*, ACM SIGKDD, SIAM SDM, IEEE ICDM, and ACM RecSys. He was the program committee member at ACM SIGKDD, IEEE ICDM, etc.

Enhong Chen (SM'07) received the Ph.D. degree from the University of Science and Technology of China (USTC). He is a professor and vice dean of the School of Computer Science, USTC. His general area of research includes data mining and machine learning, social network analysis, and recommender systems. He has published more than 100 papers in refereed conferences and journals, including the *IEEE Transactions on Knowledge and Data Engineering*, the *IEEE Transactions on Mobile Computing*, the *IEEE Transactions on Industrial Electronics*, the *ACM Transactions on Knowledge Discovery from Data*, ACM SIGKDD, IEEE ICDM, and NIPS.

He was on program committees of numerous conferences including SIGKDD, ICDM, and SDM. He received the Best Application Paper Award on KDD-2008, the Best Research Paper Award on ICDM-2011, and the Best of SDM-2015. His research is supported by the US National Science Foundation for Distinguished Young Scholars of China. He is a senior member of the IEEE.

Xi Zhang received the Ph.D. degrees from the University of Science and Technology of China (USTC), and City University of Hong Kong. He is a professor of the College of Management and Economics, Tianjin University. His general area of research includes knowledge management, social network, and Fintech. He has published more than 50 papers in refereed conferences and journals, such as Journal of Management Information Systems, IEEE Transactions on Engineering Management, Journal of Knowledge Management, etc. He was on co-chair of cognitive computing track of WWW and program committees of WWW (OBTA workshop), HICSS, etc. His research is supported by the Outstanding Young Scholars Fund of National Natural Science Foundation of China.

Tong Xu received the Ph.D. degree in University of Science and Technology of China (USTC), Hefei, China, in 2016. He is currently working as a Postdoctoral Researcher of the Anhui Province Key Laboratory of Big Data Analysis and Application, USTC. His research interests include Social Network and Media Analysis, Mobile Computing, Recommender System and other data mining related techniques.

He has authored more than 20 journal and conference papers in the fields of social network and social media analysis, including ACM SIGKDD, AAAI, IEEE ICDM, SDM, etc. Currently, he serves as a committee member of the Youth Working Committee, Chinese Information Processing Society (CIPS), and a communication committee member of the Expert Committee of Social Media Processing, CIPS.