Will I Win Your Favor? Predicting the Success of Altruistic Requests

Hsun-Ping Hsieh[1], Rui Yan[2], and Cheng-Te Li[3](✉)

[1] National Cheng Kung University, Tainan, Taiwan
sandoh714@gmail.com
[2] Baidu Inc., Beijing, China
yanrui02@baidu.com
[3] Academia Sinica, Taipei, Taiwan
ctli@citi.sinica.edu.tw

Abstract. As those in need increasingly ask for favors in online social services, having a technique to accurately predict whether their requests will be successful can instantaneously help them better formulating the requests. This paper aims to boost the accuracy of predicting the success of altruistic requests, by following the similar setting of the state-of-the-art work ADJ [1]. While ADJ has an unsatisfying prediction accuracy and requires a large set of training data, we develop a novel request success prediction model, termed *Graph-based Predictor for Request Success* (GPRS). Our GPRS model is featured by learning the correlation between success or not and the set of features extracted in the request, together with a label propagation-based optimization mechanism. Besides, in addition to the textual, social, and temporal features proposed by ADJ, we further propose three effective features, including centrality, role, and topic features, to capture how users interact in the history and how different topics affect the success of requests. Experiments conducted on the requests in the "Random Acts of Pizza" community of *Reddit.com* show GPRS can lead to around 0.81 and 0.68 AUC scores using sufficient and limited training data respectively, which significantly outperform ADJ by 0.14 and 0.08 respectively.

Keywords: Social media · Altruistic requests · Success prediction

1 Introduction

With the maturity of World Wide Web, online services provide various functions for social good. Those in need can use social communities, such as *DonorsChoose.org*, *flyingv.cc*, and *Reddit.com*, to request donations or any help. Recent studies are studying why and how a posted request can get accepted and become successful, because understanding the hidden factors driving the requests to be satisfied by givers can be a great benefit for not only those in need to write their requests but also more people to help promote such kind of requests [16,21]. Existing work had identified some key factors, including the scale of the request

J. Bailey et al. (Eds.): PAKDD 2016, Part I, LNAI 9651, pp. 177–188, 2016.
DOI: 10.1007/978-3-319-31753-3_15

(e.g., a simple question vs. a big financial need) [22], whether the giver receives in return [21], and the social interaction between the receiver and the giver [8]. Furthermore, to uncover how the linguistic factor solely influences the success of requests, the ADJ work [1] focuses on *altruistic* requests, in which the giver receive *no* rewards. Based on the qualitative analysis of linguistic factors on decision making [26], ADJ shows that the linguistic presentation, including narrative structure, politeness, evidentiality, and reciprocity, has strong correlation to the success of requests. They also demonstrate the predictability of the success of requests by treating the measured scores of these factors as feature values.

This paper aims to boost the performance of predicting the success of altruistic requests via following the similar setting and data used by ADJ [1]. While what ADJ mainly contributes is developing a series of textual features to characterize the linguistic presentation of altruistic requests, study their correlation with request success, and use the logistic regression model to test the predictability, we need to point out four aspects of insufficiency of ADJ. First, the prediction accuracy is not satisfied (the AUC score is only 0.67). Second, ADJ does not study the importance of features, but in practice different requests can resort to or concentrate on various factors to seek for the success. Third, ADJ uses a large set of requests to build the predictive model (70 % for training and 30 % for testing). However, in real-world applications, we might not have many request data with the labels of success and unsuccess for training. Fourth, the features considered in ADJ cannot model how users interact with each other affect the success, and more importantly how the topics of the requests have impact on the success. We believe user interactions and topic information also play a deterministic role in the success prediction of altruistic requests.

We think a highly accurate predictive model can bring practical advices for requesters to optimize their presentation in a *real-time* manner when asking for favors. To have a more powerful method with limited training data to accurately predict the success of altruistic requests, we devise a novel model, *Graph-based Predictor for Request Success* (termed *GPRS*). In addition to the three features (i.e., textual, social, and temporal) proposed by ADJ, we further propose three additional features, including centrality, role, and topic. Our GPRS model is designed to jointly learn the feature weights and predict the success of query altruistic requests. We evaluate the effectiveness of our GPRS model using the dataset provided by ADJ, i.e., "Random Acts of Pizza" (RAOP)[1] in *Reddit.com*, an online community established for giving away free pizza to strangers that ask for one. The results exhibit GPRS is able to not only significantly outperform ADJ with the AUC score up to 0.8, but also produce a satisfying accuracy (AUC score = 0.68) using only 20 % request data for training.

2 Dataset and Features

2.1 Data

We use the Pizza Request Dataset[2] compiled by ADJ [1], which is the entire collection of the Random Acts of Pizza Subreddit (RAOP) from December 8,

1 www.reddit.com/r/Random%5fActs%5fOf%5fPizza.
2 http://cs.stanford.edu/%7ealthoff/raop-dataset/.

2010 to September 29, 2013, in *Reddit.com*. Totally there are 5728 altruistic pizza requests (24.6 % success rate) and 1.87M relevant posts by RAOP users (for computing user features).

2.2 Features

We expand the list of features proposed by ADJ to develop a novel prediction model. We first describe these features considered in ADJ, including textual, social, and temporal. These features mainly models how language usages, social reputation, and the request time/date affect the success of requests. Then we give the details about another three of our proposed novel features, including centrality, role, and topic. Our three features are designed to further capture how user interactions in *Reddit.com* via comments and the topic information of the requests have impact on the success, and thus can be considered as the important and informative complements of the previous three features.

Textual Features. There are six categories of textual features, which are designed to capture whether or not people will help the requesters based on the textual contents of the requests.

1. **Narrative.** The narrative of a request can significantly determine the success of that request [16,21]. We follow ADJ to use the relevant vocabularies from "Linguistic Inquiry and Word Count (LIWC)" [25] to characterize five different narratives (Money, Job, Student, Family, Craving) (please refer to ADJ for the detailed vocabularies), and compute the word counts of each narrative to be the features.
2. **Politeness.** Expressing in a polite manner can leave positive impression to the potential receivers. Some qualitative studies have shown that people have higher potential to help polite requesters [6]. We take advantage of 20 politeness strategies developed in the computational politeness model [12]. We consider that each of these politeness strategies is a binary feature value that indicates whether or not such strategy is used in the request text. Stanford Dependency Parser [11], together with Regular expression and the lexicons of each politeness strategy, is used to extract the features.
3. **Evidentiality.** To make the requests more convincing, the requesters can provide images to show the evidence of what they state and their need. For example, the pictures about the house status, the screenshot of bankbook, and the proof of disabled or unemployment. By using regular expression, we compute whether or not a request contains an image URL and the number of image URLs as features.
4. **Reciprocity.** People tend to help if they received help themselves [28]. We extract the binary feature that indicates whether or not the past posts of the request contain any phrases like "pay it forward," "pay it back," or "return the favor."
5. **Sentiment.** People is more likely to help those delivering positive mood [14]. Therefore we use the Stanford CoreNLP Package[3] to estimate whether the

[3] http://nlp.stanford.edu/software/.

sentiment of a request text exceeds the average fraction of positive sentiment sentences, and compute the counts of lexicons of positive and negative vocabularies from LIWC. In addition, we also consider the emoticons as another set of sentiment features.

6. **Length.** Longer requests can reflect more efforts of the need and be more successful. We use the word count of the request as the feature.

Social Features. Existing study showed that the inter-evaluation between people have positive correlation with the success of a request, especially user status and user similarity [2]. (1) **Status**: The *karma point* of a user measured by her activity in *Reddit.com* is used as the feature. (2) **Similarity**: People tend to help those who resemble them [10]. We use intersection size and the Jaccard similarity to compute the similarity between the requester and the other users on RAOP based on the set of Subreddits of users.

Temporal Features. We measure the specific season, month, workday or weekend, weekday, day of the month, and hour of the day of the request, as well as the number of months since the beginning of RAOP community to be the temporal features.

Centrality Features. We would like to investigate how the extent of user interaction in *Reddit.com* affects the success of a request. Users in *Reddit.com* are allowed to interact with one another through commenting others' posts. We construct an directed weighted *interaction graph* to represent their interaction behaviors, in which each node is a user and each directed edge refers to a comment from the user of the comment to the recipient. Each edge is weighted by the reciprocal of the number of comments from one user to another. Lower weight values mean frequent/stronger interaction. Then to characterize how the requester interact with others, based on the constructed interaction graph, we calculate several structural **centrality** measures, including *in-degree*, *out-degree*, *clustering coefficient*, *closeness*, *betweenness*, *eigenvector*, *PageRank*, and *HITS* scores (hub and authority), as the feature values.

Role Features. We also measure the **role** of interaction of the requester among other users in terms of *network communities* (a community refers to a set of nodes that are densely connected internally and loosely connected externally in a graph). A user who is exposed to less communities can belong to the minority that needs help [13]. But a user with connections to more communities is also capable of reaching more information, and thus has a higher possibility to earn resources [23]. Therefore, we aim to quantify how the role of a requester who plays among communities affects the success of his/her altruistic request. We detect communities in the constructed interaction graph using *Louvain's algorithm* [5]. Then based on some social theories, we measure four social roles as the features.

1. **Structural Hole.** *Structural hole theory* [7] suggests that nodes act as an intermediary between groups have higher potential to access more information and earn the favor. We measure the extent of being structural holes for nodes using the number of overlapping communities [19]. Nodes overlapped by more communities tend to act as structural holes.

2. **Structural Diversity.** A person participating in fewer diverse social contexts has been validated to have lower probability to obtain his/her material and mental need [27]. We score the structural diversity by computing the number of disjoint connected components in the neighboring induced subgraph of a node. Higher scores mean higher structural diversity.
3. **Bridging Effect.** Some people act as the role of transmitting information, instead of giving or receiving help. Since the bridging nodes in a graph have been proved effective in distributing information [24], we adopt their proposed measure, *rawComm*, to quantify the bridging effect a node involves.
4. **Group Core.** Individuals with more friends who connect to each other, i.e., higher *Triangle Participation Ratio* (TPR, i.e., the number of triangles involved) and thus are closer to the core of a group, have higher potential to receive both information and help [20]. Nodes with higher TRP values tend to have higher social visibility and thus have higher potential to receive help [3]. We compute the *Triangle Participation Ratio* as the feature values.

Topic Features. We assume that the success of altruistic requests may depend on which topic that a request belongs to. Topics like bereavement are more likely to evoke public sympathy and earn the favor than other topics like unemployment. Therefore, we aim at extract the hidden topic of a request text, and treat the hidden topics as the features. Based on the textual content, we consider three kinds of manners to model the hidden topics.

1. **Bag-Of-Word** (BOW) is one of the simplest but useful features in many natural language tasks. We consider the words with NN (noun), NR (proper noun), VV (verb), VA (adjective), and AD (adverb) for the BOW features. Word counts in a document are treated as the feature values.
2. **N-gram** is to estimate the likelihood of a sentence by conditional probability. Here we use N-gram to capture whether or not there are some important terms that can determine the success of the requests. By eliminating stop words, Bi-gram and Tri-gram are considered. For BOW, Bi-gram, and Tri-gram features, we perform feature selection (dimension reduction) to identify the important and discriminative features. The technique of *Recursive Feature Elimination with Cross Validation* [17] is used to select the best number and set of features.
3. **LDA Hidden Topic.** We exploit the *Latent Dirichlet Allocation* (LDA) [4], which is a well-known topic modeling technique, to derive the hidden topic for each request. In LDA, a document-word matrix D is decomposed to a topic-hidden matrix H and a hidden-word matrix W, i.e., $D = H \times W$, in which a parameter n is used to determine the number of hidden topic categories. Since each row vector in H (denoted by H_v) can be regarded as a request document v's hidden topic, which is represented by the distribution over hidden topic categories, we use H_v as the feature values. In addition, for a request, we propose to estimate its degree of interestingness by users in terms of hidden topics. We append the user-word matrix X to D in a row-wise manner, and obtain a combined matrix M, i.e., $M = [D; X]$. LDA is applied

again for matrix decomposition: $M = H' \times W'$. We compute and sum up the Cosine similarity score $s(v)$ between the request v's hidden topic vector H'_v and each user i's hidden topic vector H'_i, i.e., $s(v) = \sum_{i \in X} cos(H'_v, H'_i)$. The obtained score $s(v)$ is treated as the feature that characterizes the degree of interestingness for the request v over all the users.

3 The Proposed GPRS Model

We devise a novel model, *Graph-based Predictor for Request Success* (**GPRS**), to predict whether or not a request will be successful (i.e., a binary success label, 0 or 1). GPRS consists of two stages: constructing a *Request Graph*, and *Propagation-based Optimization*. The basic idea lies in representing the feature similarity-based correlation between requests in a graph structure, and jointly learning the feature weights and computing the success labels of unseen requests by spreading the probabilities of success labels in the request graph such that those requests with higher similarity with each other have the same success label (0 or 1).

3.1 Constructing Request Graph

A *Request Graph* (RG) $G = (\mathcal{V}, E)$ is devised to model the feature-based correlation between request nodes, in which $\mathcal{V}$ is the set of nodes and $\mathcal{V} = V \cup U$, where V and U are the node sets of training and testing requests respectively. The construction of RG has two parts: (1) each testing request node $u \in U$ is connected to the top-k_1 similar training node $v \in V$, and (2) each testing request $u \in U$ is connected to the top-k_2 similar testing requests $u' \in U (u' \neq u)$ in terms of features, where k_1 and k_2 are determined by a parameter $\lambda \in [0, 1]$: $k_1 = \lambda \cdot |V|$ and $k_2 = \lambda \cdot |U|$. We will show how λ affects the effectiveness and efficiency our model. On the other hand, each request $x \in \mathcal{V}$ is associated with two probabilities, $P_{s_x}(x)$, corresponding to its success label $s_x = 0$ or $s_x = 1$. $P_1(x)$ and $P_0(x)$ are the probabilities that the request x is successful or not respectively. For each training request $v \in V$ whose success label is 1 (i.e., $s_v = 1$), we always fix $P_1(v) = 1$ and $P_0(v) = 0$; $P_0(v) = 0$ and $P_0(v) = 1$ if $s_v = 0$. We also initialize $P_1(u) = 0$ and $P_0(u) = 0$ for each testing request $u \in U$.

Each edge in RG is associated as a weight that represents the feature-based correlation between requests. Given a certain feature $\mathcal{F}_d$, the *feature-based request correlation* $frc_{\mathcal{F}_d}(x, y)$ between nodes x and y, $(x, y) \in E$, can be derived from their feature difference $frc_{\mathcal{F}_d}(x, y) = \varDelta\mathcal{F}_d(x, y)$, where $\varDelta\mathcal{F}_d$ is their feature difference, defined by $\varDelta\mathcal{F}_d = \|\mathbf{f_d}(x) - \mathbf{f_d}(y)\|$. Given a list of features $F = \{\mathcal{F}_d\}$ ($d = 1, ..., m$, m is the number of features), we compute *feature-aware request correlation* value $frc(x, y)$ via the weighted sum of their correlation $frc_{\mathcal{F}_d}$, given by:

$$frc(x, y) = \exp(-\sum_{d=1}^{m} \pi_d^2 \times frc_{\mathcal{F}_d}(x, y)), \tag{1}$$

where π_d is the weight of feature $\mathcal{F}_d$. The combined correlation is considered as the edge weight $w_{x,y} = frc(x, y)$ for edge $(x, y) \in E$ in RG.

3.2 Propagation-Based Optimization

The idea is two-fold. First, the probabilities of training request $v \in V$, $P_0(v)$ and $P_1(v)$, are used to infer the probabilities of testing request $u \in U$, $P_0(u)$ and $P_1(u)$. Second, $P_0(u)$ and $P_1(u)$ are inferred from both u's neighboring training and testing requests in RG, expressed by $P_s(u) = \frac{1}{deg_u} \sum_{(u,x) \in E} w_{u,x} \cdot P_s(x)$, where deg_u is the degree of node u in RG. Putting these together, we seek for an optimal set of edge weights $\mathcal{W}$ in RG, where edge weights can be further determined by feature weights π_d, such that after inference, the testing request and its neighboring requests that possess similar features tend to have close probabilities, which lead to the same success labels. When the iterative propagation process is finalized, we can choose the success label $s_u^\star$ with the higher probability to be the predicted result, given by $s_u^\star = argmax_{s_u}\{P_{s_u}(u)\}, s_u = 0, 1$.

Recall that edge weights are obtained by the weighted sum over $frc_{\mathcal{F}_d}(x, y)$ of features $\mathcal{F}_d$ with feature weights $\{\pi_d\}$. That said, the determination of feature weights first influences the edge weights, and then edges weights take effect on the inference of success labels for testing requests. Hence, our ultimate goal is to learn a set of feature weights from training and testing requests in RG.

We propose a heuristic objective for learning $\{\pi_d\}$: minimizing the *average entropy of success probabilities* $H(P^U)$ for testing requests $u \in U$:

$$H(P^U) = \frac{1}{|U|} \times - \sum_{s=0,1} P_s(u) \log P_s(u) + (1 - P_s(u)) \log(1 - P_s(u)), \tag{2}$$

where $|U|$ is the number of testing requests in RG. Our idea is that assigning the testing requests $u \in U$ the success probabilities $P_0(u)$ and $P_1(u)$ that produce the lower entropy values can make the inference be less uncertain and higher confidence. Therefore, we take advantage of the minimization of $H(P^U)$ so that through Eq. (2) the inferred success probabilities of each testing request tends to be squeezed and constrained at a success label $c_u^\star$ which possesses the highest probability.

We design a *mutually reinforced* flow to iteratively minimize $H(P^U)$ during the propagation process in RG: the learned feature weights π_d triggers an update of edge weights $w_{x,y}$ that update the success probabilities $P_1(x)$ and $P_0(x)$ for every testing request $x \in U$, which further determine their average success probability entropy $H(P^U)$ to be minimized in Eq. (2). This flow proceeds iteratively till the convergence is reached. To enable this flow, we exploit the technique of *gradient descent* on π_d to obtain an updated set of feature weights $w_{x,y}$ that minimizes $H(P^U)$. The gradient $\frac{\partial w_{x,y}}{\partial \pi_d}$ can be derived by computing $\frac{\partial H(P^U)}{\partial \pi_d}$:

$$\frac{\partial H(P^U)}{\partial \pi_d} = \frac{1}{|U|} \sum_{x \in U} \log \frac{1 - P(x)}{P(x)} \frac{\partial P(x)}{\partial \pi_d}. \tag{3}$$

Using the chain rule of differentiation, we can have the final gradient as:

$$\frac{\partial w_{x,y}}{\partial \pi_d} = 2 \cdot w_{x,y} \cdot frc_{\mathcal{F}_d}(x, y) \cdot \pi_d. \tag{4}$$

In short, in each iteration, we update the feature weights $\pi_d = \pi_d - 2 \cdot w_{x,y} \cdot frc_{\mathcal{F}_d}(x, y) \cdot \pi_d$. Then a new set of edge weights $w_{x,y}$ can be derived using the updated π_d via Eq. (1). Based on new $w_{x,y}$, we can generate the new success probabilities of testing requests $P_1(x)$ and $P_0(x)$ from x's neighbors in RG. Then the average success probability entropy $H(P^U)$ is updated accordingly via Eq. (3). The iterative updating procedure will continue and be terminated till $H(P^U)$ converges. Finally, using the derived success probabilities of each testing request $x \in U$, we can find its predicted success label $s_x^\star$. Note that we can prove the convergence by establishing a reduction from the graph-based label propagation [29]. However, due to the page limit, we skip the proof.

4 Evaluation

The experiment consists of four parts. First, we aim to show whether the proposed GPRS model can perform better than ADJ [1] using the same volume of training requests. Second, we will show GPRS can still work well using a small set of training requests. Third, we will present the effectiveness of each feature category, i.e., textual, social, temporal, centrality, role, and topic. Fourth, the time efficiency of GPRS by varying the parameter λ will be reported.

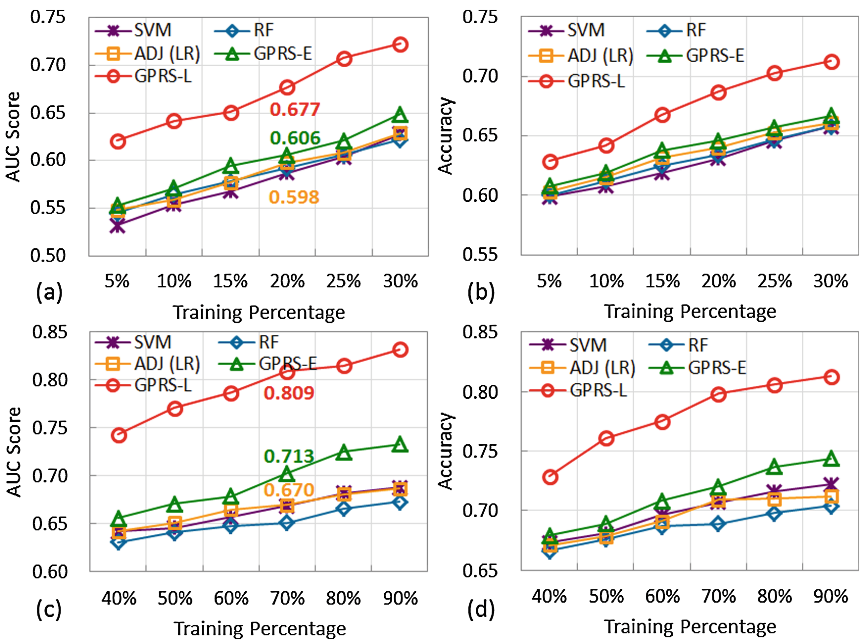

Fig. 1. AUC and accuracy by varying the percentage of training data (Color figure online).

Settings. We vary the percentage of requests for training data using two scales: 40 %–90 % and 5 %–30 % to evaluate the first two parts. We set $\lambda = 0.1$ by default, and vary λ (it determines the size of the request graph) to show the time efficiency of GPRS-W using 70 % requests as the training data. In addition, to further know whether learning feature weights π_d can benefit the prediction accuracy, we divide our GPRS into two versions: equally assigned (**GPRS-E**) and automatically learned (**GPRS-L**) from the validation set. We have two evaluation metrics. The first one follows ADJ to use the *Area Under Receiver-Operating Characteristic (ROC)* Curve (AUC). Second, we define an accuracy measure $acc = \frac{\#hits}{|U|}$, where $\#hits$ is the number of correctly predicted testing requests and $|U|$ is the number of testing requests. We randomly select the corresponding percentages of training and testing data, and repeat the experiment up to 100 times. The average AUC score and the average accuracy are reported.

Competitors. In addition to L_1-penalized logistic regression (LR) model [15] used by ADJ, we further compare our GPRS model with several typical supervised learning methods (treated as baseline models), including Support Vector Machine (SVM) [9], and Random Forest (RF) [18].

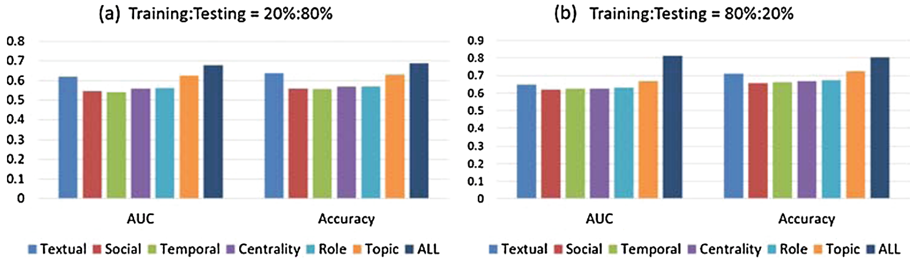

Fig. 2. AUC and accuracy of each feature category using our GPRS-L model under training:testing = 80 %:20 % and training:testing = 20 %:80 %.

The results under different training percentages are shown in Fig. 1. We can find that both GPRS-L and GPRS-E outperform ADJ and other competitors in every case, especially for GPRS-L whose AUC and accuracy scores are significantly higher than ADJ. Since ADJ uses solely 70 % training percentage to have 0.67 AUC, we especially highlight the results at 70 % for GPRS and show that both AUC and accuracy of GPRS-L are up to 0.8. It is also worthwhile to notice that GPRS-L can still achieve 0.67 AUC using only 20 % requests for training, and exhibits the prediction power of GPRS-L when tackling rare training data. Besides, the result that GPRS-L beats GPRS-E validates the efficacy of learning feature weights.

The resulting effectiveness of each category of feature using our GPRS-L model is shown in Fig. 2. We can find that our proposed three feature categories (i.e., centrality, role, and topic) lead to the competitive performance, comparing with the three feature categories proposed by ADJ (i.e., textual, social, and temporal). Among our proposed three feature categories, the topic feature can generate a bit more accurate prediction results than the textual feature, which is the best among the three features used in ADJ. Such outcome demonstrates the topic of the request can be more deterministic than the language usage (e.g., politeness and narrative), the provided evidence, and the conveyed sentiment, which are modelled by the textual feature. In addition, it is worthwhile to notice that the performance of the topic feature will be better (compared to the textual feature) if more training data are used (e.g., the performance of the topic feature under 80 % training is better than that under 20 % training data). We think the reason could be the hidden topics can be learned only if there are sufficient training data. Finally, combing all of these six feature categories can generate the best results.

The results of time efficiency (in second) and the AUC score using our GPRS-L model under different λ values are reported in Fig. 3. There is a trade-off between accuracy and efficiency. Higher λ values that add more edges in the request graph lead to higher AUC scores, but cost more time to run the GPRS-W model. Nevertheless, we find $\lambda = 0.1$ is a good choice since it balances the prediction accuracy (up to 0.8 AUC score) and run time (only 3 s), and suggest $\lambda = 0.1$ for real usages.

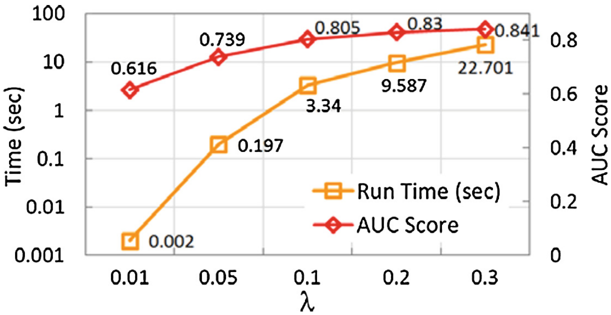

Fig. 3. AUC and run time by varying λ under the proposed GPRS-L model (Color figure online).

5 Conclusion

The contribution of our GPRS model is four-fold. First, our GPRS model is able to significantly boost the accuracy of predicting the success of altruistic requests. Second, it is capable of tacking the problem of inadequate training data while keeping the accuracy. Third, three additional features (i.e., centrality, role, and topic) are proposed and evaluated to be effective in predicting the success of requests. Fourth, the technique that jointly learns feature weights and predicts labels can be served as a novel framework to solve other NLP tasks like sentiment detection and POS tagging.

Acknowledgments. This work was sponsored by Ministry of Science and Technology of Taiwan under grant 104-2221-E-001-027-MY2. This work is also supported by Multidisciplinary Health Cloud Research Program: Technology Development and Application of Big Health Data, Academia Sinica, Taipei, Taiwan under grant MP10212-0318.

References

1. Althoff, T., Danescu-Niculescu-Mizil, C., Jurafsky, D.: How to ask for a favor: a case study on the success of altruistic requests. In: Proceedings of AAAI International Conference on Web and Social Media, ICWSM (2014)
2. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J.: Effects of user similarity in social media. In: Proceedings of the 5th ACM International Conference on Web Search and Data Mining, WSDM, pp. 703–712 (2012)
3. Backstrom, L., Huttenlocher, D., Kleinberg, J., Lan, X.: Group formation in large social networks: membership, growth, and evolution. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD, pp. 44–54 (2006)
4. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
5. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech.: Theor. Exp. **2008**(10), P1000 (2008)
6. Brown, P., Levinson, S.C.: Universals in language use: politeness phenomena. Questions and Politeness: Strategies in Social Interaction, pp. 56–311. Cambridge University Press, Cambridge (1978)
7. Burt, R.S., Holes, S.: The Social Structure of Competition. Harvard University Press, Cambridge
8. Ceyhan, S., Shi, X., Leskovec, J.: Dynamics of bidding in a P2P lending service: effects of herding and predicting loan success. In: Proceedings of the 20th International Conference on World Wide Web, WWW, pp. 547–556 (2011)
9. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 1–27 (2011)
10. Chierco, S., Rosa, C., Kayson, W.A.: Effects of location, appearance, and monetary value on altruistic behavior. Psychol. Rep. **51**(1), 199–202 (1982)
11. de Marneffe, M.-C., MacCartney, B., Manning, C.D.: Generating typed dependency parses from phrase structure parses. In: Proceedings of the Language Resources and Evaluation Conference, LREC, pp. 449–454 (2006)
12. Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., Potts, C.: A computational approach to politeness with application to social factors. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, ACL, pp. 250–259 (2013)
13. Fagan, J., Meares, T.L.: Punishment, deterrence and social control: the paradox of punishment in minority communities. Ohio State J. Crim. Law **6**, 173 (2008)
14. Forgas, J.P.: Asking nicely? The effects of mood on responding to more or less polite requests. Pers. Soc. Psychol. Bull. **24**(2), 173–185 (1998)
15. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. **33**(1), 1–22 (2010)
16. Greenberg, M.D., Pardo, B., Hariharan, K., Gerber, E.: Crowdfunding support tools: predicting success & failure. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 1815–1820 (2013)

17. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene selection for cancer classification using support vector machines. Mach. Learn. **46**(1–3), 389–422 (2002)
18. Liaw, A., Wiener, M.: Classification and regression by randomForest. R News **2**(3), 18–22 (2002)
19. Lou, T., Tang, J.: Mining structural hole spanners through information diffusion in social networks. In: Proceedings of the ACM International Conference on World Wide Web, WWW, pp. 825–836 (2013)
20. Malliaros, F.D., Vazirgiannis, M.: To stay or not to stay: modeling engagement dynamics in social graphs. In: Proceedings of ACM International Conference on Information and Knowledge Management, CIKM, pp. 469–478 (2013)
21. Mitra, T., Gilber, E.: The language that gets people to give people to give: phrases that predict success on kickstarter. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW, pp. 49–61 (2014)
22. Mollick, E.: The dynamics of crowdfunding: an exploratory study. J. Bus. Ventur. **29**(1), 1–16 (2014)
23. Pearce, J., Witten, K., Hiscock, R., Blakely, T.: Are socially disadvantaged neighbourhoods deprived of health-related community resources? Int. J. Epidemiol. **36**, 348–355 (2007)
24. Scripps, J., Tan, P.-N., Esfahanian, A.-H.: Exploration of link structure and community-based node roles in network analysis. In: Proceedings of the IEEE International Conference on Data Mining, ICDM, pp. 649–654 (2007)
25. Tausczik, Y.R., Pennebaker, J.W.: The psychological meaning of words: LIWC and computerized text analysis methods. J. Lang. Soc. Psychol. **29**(1), 24–54 (2010)
26. Teevan, J., Morris, M.R., Panovich, K.: Factors affecting response quantity, quality, and speed for questions asked via social network status messages. In: Proceedings of the AAAI International Conference on Web and Social Media, ICWSM (2011)
27. Ugander, J., Backstrom, L., Marlow, C., Kleinberg, J.: Structural diversity in social contagion. Proc. Nat. Acad. Sci. U.S.A. (PNAS) **109**(16), 5962–5966 (2012)
28. Willer, R., Flynn, F.J., Feinberg, M., Mensching, O., de Mello Ferreira, V.R., Bianchi, A.M., Choshen-Hillel, S., Weisel, O., Peng, K., Fetchenhauer, D.: Do people pay it forward? Gratitude fosters generalized reciprocity. Technical report (2013)
29. Zhou, D., Bousquet, O., Lal, T.N., Weston, J., Scholkopf, B.: Learning with local and global consistency. In: Proceedings of Advances in Neural Information Processing Systems, NIPS 2003, pp. 321–328 (2003)