Text-Aware Graph Embeddings for Donation Behavior Prediction

MeiXing Dong
University of Michigan
meixingd@umich.edu

Xueming Xu
University of Michigan
xueming@umich.edu

Rada Mihalcea
University of Michigan
mihalcea@umich.edu

Abstract

Predicting user behavior is essential for a large number of applications including recommender and dialog systems, and more broadly in domains such as healthcare, education, and economics. In this paper, we show that we can effectively predict donation behavior by using text-aware graph models, building upon graphs that connect user behaviors and their interests. Using a university donation dataset, we show that the graph representation significantly improves over learning from textual representations. Moreover, we show how incorporating implicit information inferred from text associated with the graph entities brings additional improvements. Our results demonstrate the role played by text-aware graph representations in predicting donation behavior.

1 Introduction

Understanding and predicting user behavior from their digital traces is important for many applications, such as recommender systems (Resnick and Varian, 1997), information filtering (Belkin and Croft, 1992), or dialogue agents (Mazare et al., 2018), as well as numerous behavioral interventions in healthcare, education, economics, and more. Prior research efforts have modeled user interests for predicting future behavior such as purchases (Pradel et al., 2011) or click-through likelihood (Qin et al., 2020), using signals like engagement with social media content or purchase history.

Traditional approaches to user behavior prediction use machine learning models that make use of input features in a linear fashion. These models, including the more advanced neural network architectures, assume that individual data samples are provided one at a time and independent of one another. Example user modeling approaches include using recurrent neural networks to encode the behavioral history of each user (Zhang et al., 2014) or linearly aggregating different parts of a user's background and behavior, such as their demographics and online posting patterns (Xu et al., 2020). Such approaches do not take full advantage of the relations between entities; for instance, two products in one's purchase history may be different but still be related to one another; or two users may have interests that are seemingly diverse, but which have some degree of similarity. Richer input representations that incorporate such relations can improve the performance of downstream machine learning models used to predict user behavior.

Graph models are a prominent way of representing relational information between entities. In particular, knowledge graphs have been used widely in the context of recommender systems. For example, one can construct a knowledge graph consisting of clothing brands and items and retrieve the most relevant or similar items to recommend to a user based on their most recent clothing purchase (Wang et al., 2019; Palumbo et al., 2018). Further, interactions between users and entities can also be included in the graph, such as clicks or purchases. Such a graph and its resulting node embeddings can better capture the relations between entities that arise from the aggregate behaviors of all the users.

However, these relations still only come from explicitly observed interactions like someone clicking on one entity and then also purchasing another entity, or multiple people co-clicking or co-purchasing the same entity. In many contexts, the resulting knowledge graph is sparse, as there is an absence of many co-occurring user-entity interactions due to factors such as a very large number of entities, or users having on average a very low number of interactions. As such, the learning models applied on these sparse graphs can be lacking.

In this paper, we explore user behavior prediction by using text-aware graph representations in the context of university alumni donations. We model alumni donation behavior through text and graph-based representations and evaluate our methods by predicting how likely a potential alumnus will donate to specific charitable funds. We conduct our experiments using the history of donations and university engagement newsletters of a large Midwest public university.

We start by building a graph representation of alumni and associated entities, such as academic majors, university funds, and articles in engagement newsletters. Alumni actions, such as donating to a fund or clicking on an article in an engagement newsletter, are represented as edges connecting an alumnus node with a fund or article node. Node embedding representations derived from this graph are thus capturing how different funds or engagement articles are related with respect to the alumni who donated to or clicked on them. We then use this graph to predict the likelihood of an alumnus to donate to a given charitable fund.

Specifically, our paper makes the following two main research contributions. First, we propose a graph framework to represent and predict user behavior, and show that it improves significantly over a linear representation that does not incorporate relational information. Second, we show how this graph representation can be further enriched with implicit links drawn using semantic connections between the textual information associated with the graph entities, leading to additional performance improvements in user behavior prediction. Overall, through experiments on a large alumni donations dataset, we demonstrate the effectiveness of using graph representations enhanced with implicit information for the purpose of user behavior prediction.

2 Related Work

2.1 Combining Graphs and Text

Graph models and knowledge bases are commonly used in a wide range of tasks. However, given the nature of dealing with discrete entities and relations, they can suffer from incomplete coverage or difficulty reasoning over entity relationships.

Advancements in representation learning on graphs have proven helpful in predictive tasks, such as link prediction (Wang et al., 2014), node classification (Cai et al., 2018), and node retrieval or recommendation (Zhao et al., 2015; Li et al., 2016). Many methods build embedding representations of graph nodes (Goyal and Ferrara, 2018) derived from the graph's link structure, using adjacency matrix factorization methods (Tang et al., 2015) or random walks (Grover and Leskovec, 2016).

Work has also been done towards creating text-aware graph embedding models. Methods include representing an entity through a text embedding of the entity name (Socher et al., 2013) and jointly learning embeddings for entities and words (Toutanova et al., 2015; Xiao et al., 2017).

In our work, we leverage node embedding methods to build continuous vector representations of university alumni and charitable funds, and show that they improve over text-based representations.

2.2 Predicting User Behavior

Much research has focused on predicting future user behavior based on user characteristics or prior behavior. Types of predicted behavior span a wide spectrum, including what online content someone will consume (Yin et al., 2014), what types of everyday activities someone does (Wilson and Mihalcea, 2020), and whether someone will persistent in personal improvement (Dong et al., 2021).

In the space of charitable giving, much prior work has targeted identifying factors behind why people choose to make monetary contributions. These factors include socio-demographic and personality characteristics such as age, level of education, income, agreeableness, and empathy (Bekkers, 2010; Snipes et al., 2010; Shier and Handy, 2012; Kitchen, 1992). In our context of university donations, prior work has looked at predicting how likely it is for an alumnus to donate a substantial amount of money based on their educational and professional background (Dong et al., 2020). While this shed light on signals of individual capacity and general inclination to donate, this did not look at which specific causes donors choose to give to.

There is substantially less insight into which specific charitable causes donors are likely to choose. Studies have primarily focused on giving among one or two types of charities, such as secular and religious causes (Helms and Thornton, 2012), or international and national causes (Rajan et al., 2009; Micklewright and Schnepf, 2009). These are mainly based on surveys (Breeze, 2013) asking people to recount their recent donations and describe personal dispositions such as values (Sneddon et al., 2020), empathy (Neumayr and Handy, 2019), and beliefs about the cause (Bachke et al., 2014). Most such studies are limited in the number of donors, donations, and charities observed.

In our work, we model donor behavior and donation choices using a large dataset of donations to

Entity type	Number
Alumni	5883
Funds	1644
Articles	283
Majors	251

Table 1: Statistics of entities in the alumni donation dataset.

different causes, connected with known histories of donor interactions with engagement efforts that indicate personal interests.

3 University Alumni Dataset

We conduct our experiments on a dataset of alumni information maintained by a large, public university in the Midwestern region of the United States. Each alumnus is tied to their educational history; we primarily use their major during their highest level of study at the university. The language used in the data is English.

We focus on those who have donated any amount back to their alma mater and who have also engaged with engineering alumni online newsletters, which are typically distributed by email on a regular basis. We have 2 years of newsletter content from January 2018 to March 2020, accompanied by the interaction history of alumni. The interaction history consists of when and how many times a click occurred, as well as what article was specifically clicked in the newsletter.

Likewise, we also have a history of donations that individual alumni have made to various causes at the university. Given our focus on those who have engaged with newsletters, the corresponding history of donations for these alumni span between January 2015 to June 2020. We show statistics about entities in our dataset in Table 1.

3.1 Donation Funds

At this university, alumni typically donate to funds with designated purposes. For instance, the “Engineering Student Emergency Fund" supports emergency needs related to the well-being of Engineering students. They have a title and an optional textual description of the fund’s purpose. Examples of funds and their descriptions are shown in Table 2. We see that fund descriptions can range from short and generic to lengthier and more detailed. Similarly, titles can also range in their descriptiveness of the fund’s purpose.

The set of all funds span different schools and countless initiatives. In our work, we consider only the 1644 funds (Tab. 1) that have been donated to by people who have clicked on engineering alumni engagement newsletters.

3.2 Engagement Newsletters

The university under consideration sends online newsletters to their alumni on a regular basis. These newsletters contain university news, such as student accomplishments, novel research findings, and alumni events. They consist of links to articles with an accompanying graphic and a short summary.

User actions are recorded, such as clicking on a particular article within the newsletter. Engagement with a newsletter is indicative of what alumni are interested in beyond their formal studies. For instance, a computer science graduate may primarily read articles about the solar car racing team or the university’s efforts to lower its carbon footprint, showing that this alumnus has personal interests in sustainability. This would not necessarily be apparent in their educational or employment history. Therefore, we utilize user interaction with engagement newsletters to model personal user interests. There are 283 articles in our dataset (Tab. 1), drawn from 49 total newsletters.

4 Representing Alumni and Funds

We aim to represent each alumnus primarily with their clicks. As seen in the previous section, every article linked within a newsletter has an accompanying short preview or summary that is displayed in the newsletter. Since this is what alumni initially see and what prompts their clicks, we use this text in our experiments, rather than the full article text.

4.1 Text Representation

Prior work has successfully represented entities in a graph as the average of the word vectors corresponding to its name (Socher et al., 2013). We therefore also encode our entities using word vectors. We represent an alumnus as their history of newsletter article clicks, which indicates their interests. We construct an alumnus embedding that is the averaged GloVe embedding of all newsletter article summaries that they have clicked on. We first compute an average GloVe embedding for each article snippet and then average over all of the article snippet embeddings to get the overall alumnus

Fund Name	Fund Description
Engineering Diversity, Equity, and Inclusion Initiatives	This fund helps provide a vibrant and inclusive climate, which leverages our strengths, broadens our perspectives and paves the way for innovation.
Engineering Student Emergency Fund	This expendable fund supports the emergency needs related to the health, safety and well-being of our Engineering students, especially during the current coronavirus pandemic.
Jane Doe Dance Scholarship Fund	This endowment provides scholarship support for undergraduate dance majors.

Table 2: Examples of funds and descriptions.

embedding. Similarly, we represent a fund using the average GloVe embedding of the words in the fund's name, department, and description.

4.2 Graph Representation

We construct a graph to encapsulate the connections between alumni, alumni majors, funds, and newsletter articles. Each unique alumnus, major, fund, and newsletter article are nodes in the graph. We include an edge between an alumnus and a fund if they have donated to it, weighted by the value of the total amount of donations they've given to this fund. We also connect an alumnus to a newsletter article if they have clicked on it, with the edge weighted by the number of clicks the person made. Funds included in the graph are only those associated with donations in the training set of our experiments. All newsletter clicks made by alumni are included, as was done in the text-only setting.

We then use a graph representation learning method to create embedding representations of the nodes. Specifically, we use the node2vec model proposed by (Grover and Leskovec, 2016). We also conducted experiments using LINE (Tang et al., 2015), but found that they yielded similar results, and therefore we only show results for node2vec.

4.2.1 Similarity Edges

While the explicit connections between entities through actions such as clicking and donating can contain a lot of information, there can still be additional connections made with additional info. Since it's unlikely that many alumni donate and click on exactly the same funds and articles, it may be difficult to capture all relations between them based on alumni behavior alone. For instance, two articles may contain very similar content but not have many overlapping clicks due to the sparseness of click data. Given the graph we have currently, the graph embedding model likely would not capture that the articles are similar based only on clicks. Similarly, two funds may be similar in their purpose and descriptions but have few overlapping donors, resulting in embeddings that do not capture their relevance to each other.

Graph edge type	Number
Alumni - Fund Edges	15,604
Alumni - Article Edges	20,184
Alumni - Major Edges	7,625
Fund - Fund Edges	72,136
Article - Article Edges	3,020

Table 3: Statistics of the graph derived from alumni clicks and donations, enhanced with implicit textual similarity edges.

To better capture these relations among articles and funds, respectively, we propose the addition of similarity edges. The addition of the proposed edges can add these relevance connections that we know inherently exist. This can allow the graph to encode that two funds are related even in the absence of explicit evidence, such as someone donating to both funds or two people clicking on the same article and donating to the same fund.

In preliminary experiments, we found that connecting all pairs of entities weighted by similarity results in lower performance embeddings, as well as much longer training times. We suspect this is due to adding too much noise to the representation through extraneous connections.

To minimize this, we only add edges if the similarity is above a certain threshold. We also give every such edge an equal weight of 1. For every pair of articles, we compute the cosine similarity between their average GloVe embeddings and add an edge between the corresponding nodes if their similarity is above 0.7. We do the same for every pair of funds, adding an edge if the similarity is above 0.8. We choose these thresholds empirically by looking at the distribution of similarities for all pairs of articles and funds, respectively, approximately keeping the upper quartile of similarity values. We give the numbers of different types of edges in the resulting graph in Table 3.

4.3 Analysis: Similarity between Alumni and Newsletter Articles

To gain further insights into the donor behavior graph model, we perform an analysis of the relationships between alumni and funds using their graph representations. We would expect the embeddings for alumni to be more similar to the embeddings of funds that they are more likely to donate to. This graph could then be used for querying for relevant entities. For instance, we could find the top funds that may be of interest to an alum.

To examine this, we compute the cosine similarity between pairs of alumni and funds where the alumnus has donated to the fund, and compare with pairs where the alumnus did not donate to the fund. We use node2vec embeddings based on the graph that has all similar edges incorporated. Further, we ensure that the model is not simply remembering known donations in this analysis by focusing on the subset of donations that occur in 2020 and *removing links between alumni and funds corresponding to these donations* from the graph, no matter which year the donation was made during. This way, we are looking at similarity of alumni and funds that are known to be related, but *that the model does not explicitly have knowledge about*; their similarity therefore comes solely from other alumni behavior and semantic connections. We show the distribution of similarities in Figure 1. Using a two-sided T-test, we calculate the statistical significance between the donation and non-donation samples of similarity values; we designate those with a significance level of $p < 0.1$.

Notably, we see that the GloVe-based similarities do not distinguish well between alumnus-fund pairs where a donation occurred and where a donation did not occur. In fact, the non-donation pairs actually have higher similarity than the donation pairs. This implies that it is not sufficient to use only textual semantic similarity between alumni and funds for determining donation interest.

However, we see significantly higher similarities

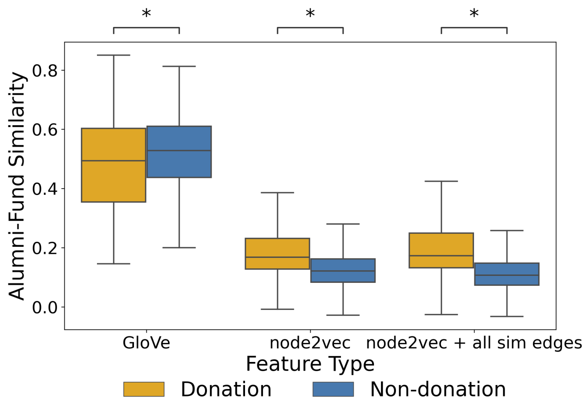

Figure 1: Distribution of similarities between pairs of alumni and funds where alumni have either donated to the fund or not. We show distributions of embedding cosine similarity based on text-only GloVe features and node2vec graph features with and without the addition of similarity edges. Statistical significance is determined using a two-sided T-test, and designated with a star (*) if $p < 0.1$.

between alumni and funds that they have donated to than between negative samples of alumni and funds when using graph embeddings. Further, this is more pronounced when similarity edges are included in the graph, yielding greater separation between pairs who have and have not donated, respectively. This shows that the graph embeddings are indeed encoding alumni behavior and interest.

5 Predicting User Behavior

We have seen that the alumni behavior graph model encapsulates relationships between entities in the resulting embedding space. We evaluate the alumni behavior graph model for downstream predictive use in the context of donation prediction. We construct a task where we predict whether an alumnus is likely to donate to a particular fund, showing that we can distinguish *which funds* someone is likely to donate to.

5.1 Experimental Setup

We focus on alumni who have both clicked on newsletter articles and have made donations. We conduct our experiments on this set of alumni, along with the funds that they have donated to, their majors (only as graph nodes), and the newsletter articles they have clicked on. We look at all pairs of alumni and the particular funds they've donated to as data samples. Donations made prior to the beginning of 2020 are considered training data and donations made in 2020 are test data. Splitting our

Alumnus-Fund Pairs	# Train	# Test
Complete	19,882	3,236
Unique	18,888	3,058

Table 4: Number of samples in the training and test sets of our task. The training samples are donations that were made prior to 2020. The test samples are donations made in 2020.

data by time reflects the real task that universities face, where we know an alumnus' history and want to predict their future donation behavior.

Funds that do not appear prior to 2020 are not included, as our graph representation models are based solely on the training data and would not be able to produce a representation for a previously unseen entity. Similarly, alumni who only appear in 2020 would be excluded from the experiments as they have no prior history and therefore would have no corresponding representation features.

We then use negative sampling to construct sample pairs where the alumnus has not donated to the fund. The training set includes an equal number of such negative samples to obtain a balanced dataset. When looking at accuracy, a balanced test set can better show the model's performance. We therefore also balance the test set. To construct a negative sample, we randomly select an alumnus and a fund from those considered in our dataset. Then, we check if the alumnus-fund pair appears as a positive sample in the corresponding data split and keep the pair if it does not appear.

Donation prediction with unique alumnus-fund pairs. We also conduct experiments in a modified setting where we predict the donation interests of alumni without knowledge of their past donations to the same funds they've donated to in 2020. We remove all alumnus-fund pairs from the training set that occur in the test set, which corresponds to removing past donations that are identical to ones in 2020. Other prior donations that alumni have made are kept. This is a more difficult task, as prior donations to a fund can be highly indicative of future donations to the same fund. Therefore, we must rely more on alumni background and the implicit relationships between different funds as well as between newsletter articles and funds.

5.2 Classification

We train a logistic regression classifier to predict whether an alumnus has donated to a given fund in 2020, based on the described data. As classification model input, we concatenate the feature representations for the alumnus and fund in a given pair. When using text-only representations, we concatenate the averaged GloVe embeddings derived from text corresponding to the alumnus and the fund in a pair, respectively (Sec. 4.1). Similarly, when we use graph-based representations, we concatenate the node2vec embeddings of the nodes corresponding to the alumnus and the fund in a pair, respectively (Sec. 4.2).

There are funds that receive thousands of donations while others receive far fewer individual donations. This can be due to the fund being very general, such as a general scholarship fund, or a popular interest, such as a sports-related fund. On the other hand, funds with more specific or niche subjects may receive fewer donations. Such large data imbalances can lead predictive models to simply memorize the most frequently occurring funds, rather than using the embedded features to make more complex connections between alumni and funds. We empirically find that less than 1% of the funds we consider have received over 200 donations. Therefore, we downsample the number of unique donations each fund has to 200 samples.

Although we implement this downsampling, there are likely still funds or types of funds that are inherently more popular. For instance, funds supporting certain sports draw many donations from alumni of diverse backgrounds. For these types of funds, the alumnus-fund fit may not be as crucial for predicting whether someone will donate; classification models are likely to capture this. Therefore, we also predict donations where we use only features representing the fund, excluding all alumni features. Comparison with this setting can show whether pairwise alumnus-fund fit is indeed useful.

6 Results

We compare the use of text-only GloVe features and graph-based node2vec features in our experiments to evaluate the benefit of our alumni behavior graph model. Further, we evaluate our graph representations both when enhanced with text similarity-based edges and without to show the effects of this adding this implicit information to the graph. We show our alumni donation interest prediction results in Table 5.

In the results, we see that the graph embedding features generally perform better than the text-only

Features	Complete Donation Pairs		Unique Donation Pairs	
	Fund Only	Alumni + Fund	Fund Only	Alumni + Fund
Text-only	0.782	*0.784*	*0.799*	0.798
Graph representations	0.781	*0.812*	*0.791*	0.778
+ article sim edges	0.774	*0.817*	*0.789*	0.778
+ fund sim edges	0.804	*0.846*	0.816	*0.824*
+ article and fund sim edges	0.798	*0.841*	0.816	*0.830*
All (GloVe + node2vec w/ all edges)	0.824	***0.856***	0.848	***0.855***

Table 5: Results from the donation behavior prediction task. Left: Training set contains the complete prior donation history of alumni in test set. Right: Donations made in 2020, in the test set, are removed from the training set. Italicized values designate the highest performance for a given feature type and experimental setting. Bold values designate the highest performance in the experimental setting overall.

features. This is in line with our hypothesis, since the text only contains information about the semantic content, but nothing about how it is related to any other entities. Further, such relations would be difficult for the machine learning model to pick up through the prediction task, as alumni generally do not individually donate to many funds and there is likely little overlap between different people. This sparsity of connections are typical in many recommendation systems contexts. Our framework of encoding user behavior into a graph could therefore be applied to other types of downstream tasks that aim to predict future behavior.

We see that adding implicit edges derived from the textual content of the funds and articles generally improves performance over only having explicit action edges that designate donations and clicks. Similarity links between articles are more helpful when we have knowledge of an alumnus' entire prior donation history.

Accuracy based on using only fund features is much higher than random, showing that the model is indeed learning trends in which types of funds, in terms of content and theme, are generally more well-received. We know the classifier isn't simply picking up on specific popular funds, since we downsampled frequently occurring funds.

Notably, when we use both features from alumni and funds, we generally see better performance, especially when using graph features and with fund edges added. This shows that the prediction model is capturing learning relationships between alumni and funds, and how compatible a given alumnus is as a potential donor for a fund.

When we use only unique donation pairs, we see that the results remain largely comparable with using complete donation pairs. However, the performance is lower than with the use of complete donation pairs when using only features derived from alumni, showing that the complete donation pairs prediction model learned more about donation trends of specific alumni whereas the unique donation pairs model has to understand more of the implicit relatedness between funds and articles.

Finally, we see that combining text-only GloVe features with graph-based node2vec features yields the highest performance. This implies that there is still use in having both the semantic content of the entities and their relational information, and that they are complementary to each other.

Qualitative Analysis

For a qualitative analysis, we use the node2vec model that includes all similarity edges, built from the training data with unique donations. We analyze how the model is able to retrieve relevant alumni and funds for a given alum.

Retrieving relevant funds. In Table 6, we show examples of funds that alumni have previously donated to and the funds that the model determined to have the highest cosine similarity. In the first example, the model retrieves funds that are related to the medical field and supporting research and education in the fields, which matches well with the alum's actual prior donations to funds supporting student scholarships and an endowed professorship. The second and third examples similarly show that the given alum's previous donations and most similar funds share common themes of aerospace engineering and natural history, respectively.

Retrieving relevant alumni. In Table 7, we show examples of click and donation activities of alumni and their highest (cosine) similarity alumni

Prior Donations	Top 3 Similar Funds (Similarity Score)
Engineering General Scholarship Fund Professorship in Rheumatology	Professorship in Gastroenterology and Hepatology Fund (0.40) Gastroenterology Nurse Education Fund (0.36) Gastroenterology Education and Research Fund (0.32)
Aerospace Engineering Support Aerospace Engineering Centennial Fund	Aerospace Engineering Junior Faculty Support Fund (0.47) Aerospace Graduate Research Excellence Fellowship (0.42) Aerospace Graduate Teaching Award and Scholarship (0.38)
Iconic Mastodons Movement Fund Majungasaurus Exhibit Fund	Mammoth Museum Exhibit Fund (0.44) Museum of Natural History Discretionary Fund (0.42) Museum of Natural History Membership (0.39)

Table 6: Prior donations made by a given alumnus the top 3 most similar funds with respect to the alum, determined by embedding cosine similarity. To preserve anonymity, we remove all names and specific details from fund titles. Text of the fund descriptions are not shown for brevity.

Alum's Prior Donations and Clicks	Nearest Alum's Donations and Clicks
F: Engineering General Scholarship Fund F: Mechanical Engineering Special Gifts Fund A: A high altitude long endurance aircraft	F: Engineering General Scholarship Fund F: Mechanical Engineering Special Gifts Fund A: Second place finish for the solar car team A: 3D printing 100 times faster with light
F: Engineering Entrepreneurship Fund F: Engineering Faculty Scholar Award A: Autonomous car preventing traffic jams A: Nobel Prize nomination for powerful laser pulse	F: Engineering Dean's Discretionary Fund A: Driverless future A: Solar car test A: Smart wearables improving elderly mobility

Table 7: Examples of the most similar alumnus for a given alum. To preserve anonymity, we do not show names and remove all identifying information within fund descriptions and article titles. We show the donations and clicks made by the alumni. F - Fund; A - Article

neighbors. In the first example, the chosen alum's donations and clicks are related to mechanical engineering. The most similar alumnus has also donated to mechanical engineering funds and clicked on mechanical engineering-related articles, which shows that nearest alumni neighbors' interests and behaviors match well with the chosen alumni. Likewise, the alumnus in the second example and their most similar alumnus both share interest in autonomous vehicles and research advancements.

7 Conclusion

In this work, we explored the use of text-aware graph representations for user behavior prediction. Using a large dataset consisting of university alumni donations and their interests as expressed through click-throughs on a university newsletter, we showed that the use of a graph framework to explicitly encode the relations between user behaviors and user interests leads to significant improvements over simple linear representations.

Moreover, we showed how further improvements can be obtained by enhancing the graph with implicit links inferred from the semantic distance between graph entities' associated textual data. Our results demonstrate the role played by graph representations using explicit and implicit relations for the prediction of user behavior.

Future work can expand upon our results and explore how textual semantic links behave with different datasets with heterogeneous graph algorithms, as well as in larger-scale data settings combined with transformer-based algorithms.

Acknowledgments

This material is based in part upon work supported by the University of Michigan Fellowship program and by the John Templeton Foundation (grant #61156). Any opinions, findings, and conclusions or recommendations expressed in this material are

those of the authors and do not necessarily reflect the views of the University of Michigan or the John Templeton Foundation.

References

Maren Elise Bachke, Frode Alfnes, and Mette Wik. 2014. Eliciting Donor Preferences. *Voluntas*, 25(2):465–486.

René Bekkers. 2010. Who gives what and when? A scenario study of intentions to give time and money. *Social Science Research*, 39(3):369–381.

Nicholas J Belkin and W Bruce Croft. 1992. Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(12):29–38.

Beth Breeze. 2013. How donors choose charities: the role of personal taste and experiences in giving decisions. *Voluntary Sector Review*, 4(2):165–183.

Hongyun Cai, Vincent W. Zheng, and Kevin Chen Chuan Chang. 2018. A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications. *IEEE Transactions on Knowledge and Data Engineering*, 30(9):1616–1637.

MeiXing Dong, Rada Mihalcea, and Dragomir Radev. 2020. Extending sparse text with induced domain-specific lexicons and embeddings: A case study on predicting donations. *Computer Speech and Language*, 59.

MeiXing Dong, Xueming Xu, Yiwei Zhang, Ian Stewart, and Rada Mihalcea. 2021. Room to Grow: Understanding Personal Characteristics Behind Self Improvement Using Social Media. In *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media*, pages 153–162. Association for Computational Linguistics (ACL).

Palash Goyal and Emilio Ferrara. 2018. Graph embedding techniques, applications, and performance: A survey. *Knowledge-Based Systems*, 151:78–94.

Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable Feature Learning for Networks. *KDD : proceedings. International Conference on Knowledge Discovery & Data Mining*, 2016:855–864.

Sara E. Helms and Jeremy P. Thornton. 2012. The influence of religiosity on charitable behavior: A COPPS investigation. *The Journal of Socio-Economics*, 41(4):373–383.

Harry Kitchen. 1992. Determinants of charitable donations in Canada: A comparison over time. *Applied Economics*, 24(7):709–713.

Juzheng Li, Jun Zhu, and Bo Zhang. 2016. Discriminative Deep Random Walk for network classification. In *54th Annual Meeting of the Association for Computational Linguistics, ACL 2016 - Long Papers*, volume 2, pages 1004–1013.

Pierre-Emmanuel Mazare, Samuel Humeau, Martin Raison, , and Antoine Bordes. 2018. Training millions of personalized dialogue agents. In *Proceedings of the Conference in Empirical Methods in Natural Language Processing*.

John Micklewright and Sylke V. Schnepf. 2009. Who gives charitable donations for overseas development? *Journal of Social Policy*, 38(2):317–341.

Michaela Neumayr and Femida Handy. 2019. Charitable Giving: What Influences Donors' Choice Among Different Causes? *Voluntas*, 30(4):783–799.

Enrico Palumbo, Giuseppe Rizzo, Raphaël Troncy, Elena Baralis, Michele Osella, and Enrico Ferro. 2018. Knowledge graph embeddings with node2vec for item recommendation. In *European Semantic Web Conference*, pages 117–120. Springer.

Bruno Pradel, Savaneary Sean, Julien Delporte, Sébastien Guérif, Céline Rouveirol, Nicolas Usunier, Françoise Fogelman-Soulié, and Frédéric Dufau-Joel. 2011. A case study in a recommender system based on purchase data. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 377–385.

Jiarui Qin, Weinan Zhang, Xin Wu, Jiarui Jin, Yuchen Fang, and Yong Yu. 2020. User behavior retrieval for click-through rate prediction. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2347–2356.

Suja S Rajan, George H Pink, and William H Dow. 2009. Sociodemographic and personality characteristics of Canadian donors contributing to international charity. *Nonprofit and Voluntary Sector Quarterly*, 38(3):413–440.

Paul Resnick and Hal R Varian. 1997. Recommender systems. *Communications of the ACM*, 40(3):56–58.

Micheal L. Shier and Femida Handy. 2012. Understanding online donor behavior: the role of donor characteristics, perceptions of the internet, website and program, and influence from social networks. *International Journal of Nonprofit and Voluntary Sector Marketing*, 17(3):219–230.

Joanne N Sneddon, Uwana Evers, and Julie A Lee. 2020. Personal Values and Choice of Charitable Cause: An Exploration of Donors' Giving Behavior. *Nonprofit and Voluntary Sector Quarterly*, 49(4):803–826.

Robin L Snipes, Sharon L Oswald, Robin L Snipes, and Sharon L Oswald. 2010. Charitable giving to not-for-profit organizations: factors affecting donations to non-profit organizations.

Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Y Ng. 2013. Reasoning with neural tensor networks for knowledge base completion. In *Advances in Neural Information Processing Systems*.

Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015. LINE: Large-scale information network embedding. In *WWW 2015 - Proceedings of the 24th International Conference on World Wide Web*.

Kristina Toutanova, Danqi Chen, Patrick Pantel, Hoifung Poon, Pallavi Choudhury, and Michael Gamon. 2015. Representing text for joint embedding of text and knowledge bases. In *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, pages 1499–1509. Association for Computational Linguistics.

Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 950–958.

Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge Graph Embedding by Translating on Hyperplanes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 28(1).

Steven R Wilson and Rada Mihalcea. 2020. Predicting human activities from user-generated content. In *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pages 2572–2582.

Han Xiao, Minlie Huang, Lian Meng, and Xiaoyan Zhu. 2017. SSP: Semantic space projection for knowledge graph embedding with text descriptions. In *31st AAAI Conference on Artificial Intelligence, AAAI 2017*, pages 3104–3110.

Zhentao Xu, Verónica Pérez-Rosas, and Rada Mihalcea. 2020. Inferring Social Media Users' Mental Health Status from Multimodal Information.

Hongzhi Yin, Bin Cui, Ling Chen, Zhiting Hu, and Zi Huang. 2014. A temporal context-aware model for user behavior modeling in social media systems. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 1543–1554.

Yuyu Zhang, Hanjun Dai, Chang Xu, Jun Feng, Taifeng Wang, Jiang Bian, Bin Wang, and Tie-Yan Liu. 2014. Sequential Click Prediction for Sponsored Search with Recurrent Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 28(1).

Yu Zhao, Zhiyuan Liu, and Maosong Sun. 2015. Representation learning for measuring entity relatedness with rich information. In *IJCAI International Joint Conference on Artificial Intelligence*, volume 2015-Janua, pages 1412–1418.