A Supervised Machine-learning Approach To Multidimensional Nonprofit Classification System: Experimentation, Validation, And Replication

Abstract

The National Taxonomy of Exempt Entities (NTEE) has been used for classifying the nonprofit organizations in the United States for several decades. However, major countries in the world do not have a classification system for the nonprofit sector. This paper achieves three major goals: 1) devising a machine learning model which can classify the nonprofits using mission statements, 2) inventing a functional classification system which can be applied to different countries, 3) test the accuracy of the model and classification system. We first created a classification system cross different countries by matching existing standards, then compiled the training and testing datasets for China (data from China Foundation Center and Research Infrastructure of Chinese Foundations), United Kingdom (data from ****), and United States (data from National Center for Charitable Statistics and Internal Revenue Service). We finally test the accuracy of major text classification algorithms using country-specific training datasets and a pooled dataset. Implications and limitations are discussed.

1 Introduction

Although the voluntary and philanthropic organizations have long been existent for numerous centuries, the so-called "nonprofit sector" was only coined in the 1970s by scholars, policy makers, and nonprofit practitioners (Hall, 2006). A major reason for assembling the diverse organizations as a conceptual whole is to legitimize the existence of these organizations and the benefits these organizations receive (Barman, 2013; Hall, 2006). From Durkheim's 2012 perspective, the order and structure of a society can be reflected by a classification system. The National Taxonomy of Exempt Entities (NTEE) developed by the National Center for Charitable Statistics (NCCS), the most widely used classification system, is one of the efforts legitimizing the existence of nonprofit sector (Hodgkinson, 1990; Hodgkinson & Toppe, 1991). As Barman (2013, p. 105) cite Clarke and Casper (1996, p. 601): "The ways in which different entities (people, animals, plants, diseases, etc.) are organized into classificatory groups reveal something of the social, cultural, symbolic, and political contexts within which classifications occur."

Specifically, a classification system like NTEE has been used for many practical and academic purposes. For example, it provides a framework on which the social and economic activities of nonprofits can be mapped and compared with other types of organizations in a society (e.g., Roeger, Blackwood, & Pettijohn, 2015). Scholars also use NTEE codes for sampling purposes (e.g., Carman & Fredericks, 2010; Okten & Weisbrod, 2000) or as independent variables (Sloan, 2009). The invention of an international classification system, although challenging, can be the cornerstone for studying "global civil society" (Lester M. Salamon & Anheier, 1992; Lester M Salamon, Anheir, & coaut, 1996; Vakil, 1997).

The NTEE classification system, although the best we have, still has several major drawbacks. First, because it only assign one major category code to an organization, it cannot accurately describe a nonprofit organization's programs which are usually diverse and across several domains (Grønbjerg, 1994, p. 303). Although another classification system assigning purpose codes to programs was developed (Lampkin, Romeo, & Finnin, 2001), it is not widely used (why?). Second, the assignment of NTEE codes is not complete because it is "based on an assessment of program descriptions contained in Parts 3 and 8 of the Form 990" and "program descriptions were only available for some organizations" (National Center for Charitable Statistics, 2006, p. 16). A recent study found the number of organizations in Washington State with a specific NTEE code could be significantly increased if the mission statements were used for coding (Fyall, Moore, & Gugerty, 2018). Third, NTEE codes are static but nonprofit organizations' activities may change over time. Recoding existent NTEE assignments is laborious and almost impossible for human.

By using supervised machine-learning techniques, this study advanced the NTEE classification system from the following perspective: 1) A series of datasets for training models was developed, 2) the accuracy and efficiency of popular text-classification algorithms were compared, 3) existent empirical studies were replicated to test the validity of our machine-learning approach. According to the results of experimentation, validation, and replication, the combination of ABC algorithm and trained datasets produced the best results.

2 Method

2.1 Working with Texts: Dictionary, Supervised and Unsupervised Machine-Learning

The classification of texts is a typical task of automatic content analysis, and three types of methods are common to this task: dictionary, supervised, and unsupervised methods (Grimmer & Stewart, 2013, pp. 268–269). The dictionary methods use a predefined dictionary of words to classifying the texts. Although accurate, this approach is not capable to deal with the variations and contexts of language. An improved solution is to use supervised methods which are computer algorithms that can "learn" the linguistic patters in a dataset classified by human coders. Unlike the dictionary and supervised methods which require predefined categories of interest, unsupervised methods can discover linguistic patters in texts without inputting any knowledge of classification. However, the validity of unsupervised methods is a serious challenge because the classifications returned may not be theoretically meaningful. This study employs supervised methods to make the use of existing classifications and human-coded records and deal with linguistic variations and contexts.

2.2 Datasets

There are two types of datasets for supervised text classification: training dataset and testing dataset. Both datasets are collections of text records that have been classified by human coders. The machine-learning algorithms can "learn" the linguistic patters from the training dataset and then classify the records in testing dataset using the patters learned. The results generated by algorithms can be compared with those coded by human coders, and ultimate goal is to use trained models to replace human. The quality of training dataset is decisive because it must be a representative sample of the whole corpus. The training dataset can be generated by randomly sampling the whole corpus, but a better strategy is proportional sampling according to the distribution of classification scheme (Grimmer & Stewart, 2013, p. 278).

Figure **??** shows the proportions of NTEE major categories from 1989 to 2015 ...

2.3 Machine-Learning Algorithms

This study uses three common supervised machine-learning classification algorithms: Naïve Bayes, Random Forest, and Neural Network. Other than these three individual methods, an ensemble approach that combines all the three models is also experimented.

3 Results

3.1 Validity of algorithms

Use *confusion matrix* see Grimmer and Stewart (2013, p. 279).

3.2 Applying the International Nonprofit Classification System: Descriptive analysis

3.3 Applying the International Nonprofit Classification System: Replication of empirical studies

References

Barman, E. (2013). Classificatory Struggles in the Nonprofit Sector: The Formation of the National Taxonomy of Exempt Entities, 1969—1987. *Social Science History*, *37*(1), 103–141. 00021.

Carman, J. G., & Fredericks, K. A. (2010). Evaluation Capacity and Nonprofit Organizations: Is the Glass Half-Empty or Half-Full? *American Journal of Evaluation*, *31*(1), 84–104. 00107. doi:10.1177/1098214009352361

Clarke, A. E., & Casper, M. J. (1996). From Simple Technology to Complex Arena: Classification of Pap Smears, 1917-90. *Medical Anthropology Quarterly*, *10*(4), 601–623. 00057.

Durkheim, É. (2012). *The Elementary Forms of the Religious Life*. Courier Corporation.

Fyall, R., Moore, M. K., & Gugerty, M. K. (2018). Beyond NTEE Codes: Opportunities to Understand Nonprofit Activity Through Mission Statement Content Coding. *Nonprofit and Voluntary Sector Quarterly*, *47*(4), 677–701. 00000. doi:10.1177/0899764018768019

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*(3), 267–297. doi:10.1093/pan/mps028

Grønbjerg, K. A. (1994). Using NTEE to classify non-profit organisations: An assessment of human service and regional applications. *Voluntas*, *5*(3), 301–328. 00055. doi:10.1007/BF02354038

Hall, P. D. (2006). A Historical Overview of Philanthropy, Voluntary Associations, and Nonprofit Organizations in the United States, 1600–2000. In W. W. Powell & R. Steinberg

(Eds.), *The nonprofit sector: A research handbook* (pp. 32–65). 00000. Yale University Press.

Hodgkinson, V. A. (1990). Mapping the non-profit sector in the United States: Implications for research. *Voluntas*, *1*(2), 6–32. 00043. doi:10.1007/BF01397436

Hodgkinson, V. A., & Toppe, C. (1991). A new research and planning tool for managers: The national taxonomy of exempt entities. *Nonprofit Management and Leadership*, *1*(4), 403–414. 00028. doi:10.1002/nml.4130010410

Lampkin, L., Romeo, S., & Finnin, E. (2001). Introducing the Nonprofit Program Classification System: The Taxonomy We've Been Waiting for , Introducing the Nonprofit Program Classification System: The Taxonomy We've Been Waiting for. *Nonprofit and Voluntary Sector Quarterly*, *30*(4), 781–793. 00022. doi:10.1177/0899764001304009

National Center for Charitable Statistics. (2006). *Guide to Using NCCS Data*. Urban Institute. Washington, DC.

Okten, C., & Weisbrod, B. A. (2000). Determinants of donations in private nonprofit markets. *Journal of Public Economics*, *75*(2), 255–272. 00466. doi:10.1016/S0047-2727(99)00066-3

Roeger, K. L., Blackwood, A. S., & Pettijohn, S. L. (2015). The Nonprofit Sector and Its Place in the National Economy. In J. S. Ott & L. A. Dicke (Eds.), *The Nature of the Nonprofit Sector* (Third edition, pp. 22–37). 00204. Boulder, CO: Westview Press.

Salamon, L. M. [Lester M.], & Anheier, H. K. (1992). In search of the non-profit sector II: The problem of classification. *Voluntas*, *3*(3), 267–309. 00332. doi:10.1007/BF01397460

Salamon, L. M. [Lester M], Anheir, H. K., & coaut. (1996). *The international classification of nonprofit organizations ICNPO-Revision 1, 1996*. 00230 OCLC: 760476834. Baltimore, Md: The Johns Hopkins University Institute for Policy Studies.

Sloan, M. F. (2009). The Effects of Nonprofit Accountability Ratings on Donor Behavior. *Nonprofit and Voluntary Sector Quarterly*, *38*(2), 220–236. 00136. doi:10.1177/0899764008316470

Vakil, A. C. (1997). Confronting the classification problem: Toward a taxonomy of NGOs. *World Development*, *25*(12), 2057–2070. 00559. doi:10.1016/S0305-750X(97)00098-3