

Measuring Mission Deviation in California Non-Profit Hospitals

Stanford CS224N Custom Project. Project Mentor: Swastika Dutta

Pranay Agrawal
Department of Computer Science
pranagra@stanford.edu

Nova Josephine Bradford
Department of Health Policy
novab@stanford.edu

Cesar Portocarrero
Department of CEE
cportoca@stanford.edu

Abstract

There is intense debate among nonprofit scholars as to whether revenue diversification will enhance or contaminate a nonprofit organization’s commitment to its mission. Researchers have been unable to definitively model the relationship between commercialization and nonprofit mission deviation largely due to a lack of coherent measurement paradigms. In this project, we propose a novel, NLP-based metric to measure mission deviation in nonprofit hospitals: lexical dissimilarity between hospital mission statements and annual reports. The lexical dissimilarity was determined by measuring the Euclidean distance in textual vector embeddings of annual reports and mission statements. Embeddings of size 786 were computed using three strategies (BERT, SBERT, and a novel strategy, hosBERT). We validate this construct through OLS models using theoretically-linked predictors, including financial ratios. The hosBERT model showcased the largest R^2 value (0.332), followed by SBERT (0.251) and BERT (.009). Hospitals predicted to be high in mission deviation based on financial metrics showed greater Euclidean distance than those predicted to be low in mission deviation in the middle sections of their annual reports, which suggests that Euclidean distance between annual reports and mission statements could be a valid measure of mission deviation.

1 Introduction

Since the end of the 1970s, there has been a steady trend of increasing commercialization of human services in the United States—dubbed a “creeping revolution” by Grønbjerg (2001). Commercialization is one of the primary strategies through which nonprofits can diversify their revenue streams. Proponents of nonprofit revenue diversification have argued that this practice can enable human services organizations to achieve self-sufficiency, whereas opponents argue that diversification may introduce a contradiction between an organization’s profit-seeking behavior and its moral adherence to social goals (Guo, 2006). This controversy centers on the fundamental question of whether revenue diversification will enhance or contaminate a nonprofit organization’s commitment to its mission.

This problem is especially germane to nonprofit hospitals, which make up approximately 60% of all community hospitals in the United States (Bradley Herring, 2018). Nonprofit organizations are tax-exempt, meaning that they are not required to pay federal or state taxes on their revenue. The fundamental assumption of the nonprofit model is that these organizations provide sufficient community benefit in exchange for tax-exempt status, formally known as 501(c)(3) organizations according to the US Internal Revenue Code (IRC). However, policymakers have increasingly voiced concerns that nonprofit hospitals may not be providing adequate community benefits to justify tax exemption (Bradley Herring, 2018). Given that healthcare composes approximately one-fifth of the United States GDP (over 4.3 trillion dollars), this is an extremely pressing policy issue.

According to the IRC, any organization that deviates substantially from its original objective is subjected to revocation of its 501(c)(3) status. Researchers, however, have been unable to definitively model the relationship between commercialization and nonprofit mission deviation largely due

to a lack of coherent measurement paradigms. Qualitative mission deviation has been measured through qualitative interviews with nonprofit employees and has been used as a theoretical variable in nonempirical scholarship. Unfortunately, these measurement paradigms do not allow for robust, generalizable statements about mission deviation and consequently have failed to provide actionable guidance to nonprofits. For the current project, we propose a novel measurement paradigm for mission deviation: lexical dissimilarity between nonprofit annual reports and mission statements.

2 Related Work

2.1 Mission deviation and its correlation to financial metrics

The importance of mission deviation (also referred to as “mission drift” in the literature) in non-profits is evidenced by the number of research papers related to its prevention and mitigation strategies (Cornforth, 2014; Ramus and Vaccaro, 2014; Nkabinde and Mamabolo, 2022). There is no consensus on a systematic approach to measuring mission deviation, and several approaches have been proposed in recent years. For example, a rudimentary quantitative approach was proposed by Cetindamar and Ozkazanc-Pan (2017) in the context of venture capital (VC) firms. By analyzing the frequency of a curated list of keywords in mission statements and annual reports, experts determined the presence of financial and social logic in these documents. Subsequently, a score of 0 to 1 was assigned to both categories (social and financial); if a score of 1 is not present in both categories in the VC documents, then mission drift is considered to be present in the corresponding institution. Though this represented a novel, systematic approach to assess mission drift, scoring is still a human-driven process in which experts analyze physical, syntactical, categorical, propositional, and thematic distinctions in the text.

Since the passage of the Pension Protection Act in 2006, all 501(c)(3) organizations (one category of nonprofits) are required to annually file financial information with the IRS. The IRS is required to revoke the tax exemption of any organization that fails to file these documents for three consecutive years or if said organization shows evidence of mission deviation based on its documentation (Lampkin, 2010). The impact of revocation is considerable—donors cannot deduct their contributions, grantmakers and funders will not commit funds, and the nonprofit will have to pay federal income tax. Consequently, researchers have explored the correlation between financial information and mission deviation. In a case study on Habitat for Humanity, Berrett (2021) showed that an increase in administrative expenses leads to greater effectiveness and less mission deviation, demonstrated by an increase in the number of houses built and revenue raised, though it is worth noting that these results are generally inconsistent with nonprofit literature documenting a positive association between mission deviation and administrative expenses. Additionally, results demonstrated that indicators of organizational capacity (e.g., investments into salaries, professional fees, marketing, information technology, and fundraising) mediate the link between administrative costs and effectiveness. Consequently, these organizational capacity indicators are considered proxies of mission deviation and serve this role in our current analysis.

2.2 NLP applications in organizational studies

Lexical similarity has become an increasingly popular metric in policy and organizational studies as computational methods continue to gain prominence. For instance, Mandal et al. (2017) used cosine similarity between embeddings of court documents to explore the similarity of legal cases. In this study, the authors explored the use of four embedding strategies: TF-IDF Vectorizer, Topic models (using Latent Dirichlet Allocation), word embeddings (word2vec), and Document embeddings (doc2vec) on a subset of the Indian Supreme Court case documents. The results indicate that the neural network approaches (word2vec and doc2vec) outperformed the other methodologies. The evaluation was carried out using the Pearson correlation coefficient to quantify how well the four approaches correlate to the similarity scores provided by the experts (gold standard). The best performance was obtained by the Doc2vec embedding using the whole document, with a score of 0.69.

In recent years, there has been an increasing use of transformer-based language models in policy and organizational studies. For instance, Cheng Ren (2022) experimented with seven NLP algorithms to predict whether nonprofits in their sample served immigrant populations—an important task because accurately classifying nonprofits is often necessary for researchers and policymakers alike. The models used in the study included bag-of-words (BOW), term frequency-inverse document frequency

(TF-IDF), word2vec trained on a corpus of organizational names (W2V); word2vec trained on a corpus of Google News stories (G-News), global vectors for word representation (GloVe), long short-term memory (LSTM), and bidirectional encoder representations from transformers (BERT). Of the techniques used in their study, BERT yielded the best performance, with an accuracy of .89. This model outperformed two nonmachine methods used in existing research—identification of organizations via National Taxonomy of Exempt Entities codes or keyword searches of nonprofit names.

In the current study, we aim to progress these lines of research by applying transformer-based embeddings to nonprofit annual reports and mission statements to measure mission deviation.

3 Approach

Document embedding and comparison for both mission statements and annual reports were carried out using three different strategies: full document comparison using SBERT, full document comparison using BERT (Devlin et al., 2018), and full document comparison using hosBERT, a BERT model that is pre-trained using non-profit hospital data collected for this study. We also analyze how an annual report’s deviation from its mission statement varies progressively throughout the report, referring to this approach as “progressive document comparison”. While the embeddings could have been created with any of the three embedding strategies discussed above, we present results with SBERT as the embedding strategy. It is worth noting that we saw similar results with hosBERT.

3.1 Full document comparison using SBERT

SBERT is a modification of the pre-trained BERT network implemented by Reimers and Gurevych (2019). SBERT relies on siamese and triplet network structures to generate semantically meaningful sentence embeddings which can be compared using any distance metric such as cosine similarity or Euclidean distance. In our case, we used Euclidean distance. The specific model used for the embeddings is *all-mpnet-base-v2*, which is pre-trained using SNLI, MultiNLI, and TriviaQA data to create general-purpose sentence embeddings. Each token embedding is then pooled together to output a 768-dimensional embedding.

In this first approach, the architecture presented in Figure 2a (in Appendix) was used. This procedure first divides all annual reports into 50 partitions that are then independently embedded using SBERT. The embeddings are subsequently aggregated using average pooling to obtain the final document embedding (u) as shown in Equation 1. Simultaneously, mission statements are embedded with SBERT (v) and are then compared with the annual report vector using euclidean distance, which is established as a proxy for mission deviation in the context of this study. It is worth noting that the division into 50 partitions was incorporated due to the annual reports spanning thousands of words, which exceeds SBERT’s input limits.

$$(1) \quad u = \frac{1}{50} \sum_{j=0}^{49} SBERT(annual_report_partition[j])$$

3.2 Full document comparison using BERT and hosBERT

We use the BERT model *bert-base-uncased* as a baseline to perform a full document comparison similar to SBERT. BERT is unlikely to perform well in comparison to SBERT as SBERT has a much better sentence-level understanding due to its network described above.

hosBERT performs integrated training on the original BERT implementation using annual reports and mission statements from non-profit hospital data as summarized in Figure 2b. By continuing the pre-training process on the pre-trained BERT weights, using the MLM objective by masking out 15% of tokens, we make BERT more in line with our financial dataset. With the lexical and semantical knowledge acquired during pre-training, a full document comparison approach similar to the one introduced in Figure 2a is executed. This approach was formulated taking into consideration that the BERT and SBERT embedding model is not specialized in this particular domain of knowledge and thus might not capture adequately the information presented in the documentation.

3.3 Progressive document comparison

As annual reports incorporate relatively general introduction and conclusion sections in their contents that do not contribute semantically to the text, an alternative approach that allows for progressive mission deviation estimation throughout the text on a section-by-section basis was formulated. This approach is displayed in Figure 2c. In contrast with the previous method, there is no average pooling layer to generate a single annual report embedding. Instead, there are partial embeddings (u_1, u_2, \dots, u_n) that are generated using average pooling with the neighboring k precedent and subsequent partitions as shown in Equation 2. Then, each of these partial embeddings (u_i) is contrasted with the mission statement embedding to obtain multiple euclidean distance values for the same annual report. It is worth noting that the average pooling strategy adopted in this approach with the nearest k neighbors was implemented in order to capture more contextual information presented both before and after the current analyzed partition.

$$(2) \quad u_i = \frac{1}{2k+1} \sum_{-k \leq j \leq k} SBERT(annual_report_partition[i+j])$$

4 Experiments

4.1 Data

Data was compiled from three sources: 1) community benefit reports (CBR) from California nonprofit hospitals, 2) nonprofit financial data from the Internal Revenue Service (IRS), and 3) hospital records data from the California Department of Health and Human Services (CalHHS).

Nonprofit hospitals in California are required to annually file a written report documenting their community benefits in order to retain their tax-exempt status. This includes a description of their annual activities and strategic goals. These annual reports were downloaded in PDF form for every nonprofit hospital in California. For the 2020 fiscal year (the last year for which IRS data is available), 191 hospitals filed CBPs. However, many hospitals are owned and operated by the same parent organization—for instance, Sutter Health operates 24 California hospitals under the same parent organization. As such, the total number of health organizations included in this study was 91. The PDFs were named only using words, not identification codes; for instance, a PDF would be labeled "Adventist_Health_Reedley_Community_benefit_Report_2021.pdf." In order to match the annual reports to the IRS data, we utilized three databases from CalHHS: 1) a database that lists facility names and facility numbers, 2) a database that lists facility numbers and Medicare provider numbers, and 3) a database that lists Medicare provider numbers and employer identification numbers (EIN). This allowed us to construct dictionaries to map the annual report names to EINs, thus enabling us to match annual reports to IRS data (described below).

All nonprofits in the United States are required to submit annual financial paperwork to the IRS justifying their tax-exempt status, known as Form 990. This documents all sources of revenue and all expenses, as well as a brief description of the nonprofit's mission. Form 990 data are made available in a machine-readable format directly from the IRS. Data processing was computationally-intensive because the IRS provides unlabelled XML files for every organization in the United States (over 500,000 per year). In order to select the relevant Form 990 for each of the organizations in our study, it was necessary to parse every XML file and select the ones whose employer identification number (EIN) tag with the set of EINs for California nonprofit hospitals (90 organizations from the 500,000).

On an annual basis, CalHHS collects data from individual hospitals and hospital systems including detailed facility-level reporting on services capacity, inpatient/outpatient utilization, patients, revenues and expenses by type and payer, balance sheet, and income statement. To form the complete set of data for this study, PDF CBPs were processed with a PDF reader, Form 990 XML files were parsed, and CalHHS data were extracted. Hospital-level data were matched from these three sources using EIN, hospital facility number, and Medicare provider numbers. After compiling all data sources, our final sample size was 109.

4.2 Evaluation method

Since this is an unsupervised method, there are no “right” or “wrong” results per se. However, we assess the quality of this measurement paradigm through triangulation with related measures. Since there is no existing, scalable measurement paradigm to compare our results to, we will triangulate our measure with a series of alternative metrics which, according to organizational theory, should correlate with mission deviation. As such, several financial ratios computed from the IRS financial data are broadly indicative of an organization’s current financial health. These ratios and their interpretations include the following: 1) Administrative Expense Ratio: The percentage of a hospital’s expenses that are being allocated to administrative (overhead) costs. 2) Program Expense Ratio: The percentage of expenses that a nonprofit hospital is spending on its core mission. 3) Net Margin Ratio: A hospital’s ability to operate at a financial surplus rather than a financial deficit. 4) Leverage Ratio: The degree to which a hospital’s operations are funded by debt (borrowed capital).

These ratios, along with CalHHS and IRS data, were used as features in regression models, and mission deviation (Euclidean distance between annual report and mission statement vector embeddings) was used as the outcome. Regression analyses are described in more detail below.

Exploratory data analysis included visual inspection of correlation matrices. The results of exploratory data analysis were used to select six meaningful predictor variables to include in the linear regression: the four of the hypothesized ratios described above, as well as two financial variables: Other Revenue (OR) and Executive Compensation from Related Organization (ECRO).

OR is income that cannot be categorized as one of three categories (grants, program service revenue, and investment income). Organizational scholars have found that organizations reporting “other” revenue sources may have greater commercial revenue, more sophisticated financing schemes, and a more professional staff. A positive coefficient suggests that greater “other” revenue is associated with greater mission deviation.

ECRO must be reported when executive leaders are compensated more than \$10,000 by a “parent, subsidiary, or sister organization under common control.” For example, if a hospital CEO is paid by her hospital and also by a foundation that the hospital owns, the foundation pay must be reported as ECRO. This is likely an indicator of excessive compensation, one of the most salient features of inefficient or wasteful spending in the nonprofit sector Barragato (2002). A positive coefficient indicates that greater ECRO is associated with greater mission deviation.

4.3 Experimental details

After computing embeddings as described above, we employed Euclidean distance to obtain the predicted dissimilarity scores between an organization’s mission statement and its corresponding annual report. To evaluate the effectiveness of this approach in predicting mission deviation, we fit a linear regression model (dataset size 109) between the mission deviation scores and the final ratios indicated above. Linear regression was selected due to its interpretability since our goal in the current study is to identify individual predictors of mission deviation and quantify the strength of this association. Three regression models were fit to compare the performance of our three embedding strategies. BERT is considered our baseline, SBERT is an existing model to outperform, and hosBERT is our novel pre-trained model. The effectiveness of these approaches was assessed using R-square values and by qualitatively examining the fitted line. The results of our regression analysis are discussed in the results subsection.

The hosBERT model was trained on annual reports and mission statements from 60 organizations that were not used in the linear regression. Our dataset included 110,000 sentences, with the longest sentence containing 25 words. We began by tokenizing each sentence using the pre-trained *bert-base-uncased* tokenizer. We then masked out 15% of the tokens for the MLM objective. The model was trained for 40 epochs on batch sizes of 16, using the Adam optimizer with a learning rate of 1e-4 and momentum of 0.9. While we would have preferred to try larger batch sizes, our computational resources were exceeded.

For timestep-based analysis, we perform an average pooling on the embedding for each annual report partition with a window size of 2 ($k = 2$). Next, we compute deviation values for each partition of each annual report using Euclidean distance, resulting in a vector of size 50 for each annual report that is normalized. We then split the data into two groups based on the OR financial metric, selecting

the top 25% and bottom 25% of data for each group. We selected OR for this task since our central hypothesis concerned hospital revenue diversification. This enables us to determine the average textual dissimilarity for annual reports that have a high mission deviation score based on the financial metric, as well as for those with a low mission deviation score. We then present a chart displaying the mean and standard deviation for these two groups, allowing us to analyze the results.

4.4 Results

Of the three embedding strategies, hosBERT yielded the highest R^2 value (0.332) followed by SBERT (0.251) and BERT (.009). Results from the hosBERT model are summarized in Table 1.

In spite of the noteworthy model R^2 , none of the predictors were statistically significant. In this way, our model is performing worse than expected, since we anticipated statistically significant predictor coefficients. However, a comparison between low and high-deviation hospitals based on financial metrics does suggest a relationship between our predictor and response variables. The non-significant p-values are potentially the result of the sample size of the current study resulting in insufficient statistical power to identify a significant result. The R^2 value of hosBERT indicates that we can explain approximately $\frac{1}{3}$ of the variation in mission deviation using these six predictors, which is noteworthy. As such, since this study includes all nonprofit hospital systems in the state of California, additional analyses which combine data from multiple states could be conducted in the future.

Table 1: Linear Regression Using hosBERT Embeddings

Variable	Coefficient	Std. Error	t-Statistic	p-Value
Intercept	-78.7884	59.216	-1.331	0.190
Other Revenue (Current Year)	1.665e-08	1.41e-08	1.182	0.224
Executive Compensation from Associated Organizations	1.757e-10	7.95e-09	.011	0.982
Administrative Expense Ratio	85.2451	59.689	1.430	0.160
Program Expense Ratio	85.9845	59.215	1.452	0.154
Net Margin Ratio	81.0028	59.048	1.372	0.177
Leverage Ratio	-0.4083	0.462	-0.884	0.382

Table 1: Linear regression to estimate relation between Euclidean distance (using hosBERT) and six independent financial metric variables. An R^2 value of 0.332 was observed with 109 data points.

Our comparison between hospitals with high and low mission deviation demonstrated the largest differences in the center of the document as shown in Figure 1. This is consistent with our hypotheses since annual reports begin with an introduction and end with a conclusion, so the middle section includes a description of their initiatives over the past year. This suggests that meaningful financial differences exist between high and low deviation hospitals and that the middle sections of annual reports are most likely to demonstrate salient deviations.

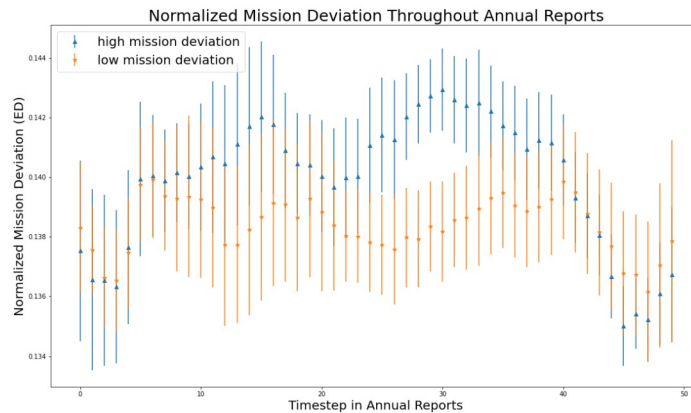


Figure 1: Errorbar plot with mean and standard deviation for normalized annual report deviation for hospitals with top 25% (blue) and bottom 25% (orange) of financial metrics determined by “Other Revenue”. Reports are split into 50 partitions to visualize deviation throughout annual reports.

5 Analysis

To qualitatively evaluate our mission deviation metric (Euclidean distance between annual report and mission statement embeddings), we compared the two hospitals which yielded the least and greatest mission deviation. The hospital with the least mission deviation was Community Hospital of the Monterey Peninsula (CHMP), a hospital in Monterey, California that was originally established to specialize in the treatment of metabolic disorders but has since grown to provide a full suite of specialty services. The hospital with the greatest mission deviation was St. Agnes Medical Center (SAMC), which owns and operates a 436-bed acute-care hospital and a home healthcare agency in Fresno, California.

Upon reviewing the mission statements and annual reports of these two hospitals, we found several noteworthy patterns that are generally consistent with our hypotheses. Notably, the mission statements of these hospitals differed significantly. CHMP states that they are “dedicated to identifying and meeting the changing healthcare needs of the people of the Monterey Peninsula and surrounding communities.” According to the mission statement of SAMC, the purpose of the organization is to “serve together in the spirit of the Gospel as a compassionate and transforming healing presence within our communities.” These illustrative examples point to a potential source of the underlying variation in our studies—organizational scholars have noted that religious hospitals face a conflict between economic and social action given that the modern industrial landscape expects hospitals to deliver high-quality medical performance rather than emotional or spiritual assistance Ann-Kathrin Seemann (2015).

A qualitative examination of the annual reports suggests meaningful differences between these organizations as well. CHMP’s annual report clearly outlines its community benefit initiatives and the value of these initiatives based on measurable, predefined objectives, such as cost-effectiveness. Examples of such initiatives include health education for patients’ families and grants to provide healthcare to uninsured patients. SAMC, on the other hand, reports a variety of “non-quantifiable benefits,” including executive leaders sitting on boards of other organizations and hospital staff coordinating volunteer programs and Christmas toy donation drives. However, even though SAMC’s annual report and mission statement is very different from CHMP, SAMC’s annual report did not capture the same textual context as its mission statement. As a result, we anticipated significant dissimilarity between the textual understanding of the mission statement and the annual report, consistent with our model’s textual embeddings.

Our qualitative examination indicated that the embeddings reflected the lexical structure of the documents accurately, consistent with the opinions of human readers. Furthermore, low mission deviation scores for CHMP corresponded with its generally good financial performance. However, the high mission deviation score for SAMC did not correspond with its financial performance, likely due to the hospital’s religious focus.

We also found that the model’s predictions of deviation compared favorably to those of human readers, especially when comparing annual reports progressively. We observed that most reports began with a statement of the mission statement, followed by a brief overview of the hospital’s vision, and then a statement from the CEO. The conclusion often summarized the hospital’s goals and aims. As a result, the beginning and end of most reports had low deviation, which was in line with qualitative expectations.

The middle of the report generally described the hospital’s strategic initiatives and the number of people served during the year. As this did not have to do with the high-level goals of the organization and involved more specific details, the words used to describe various programs may differ from the hospital’s mission statement even though these initiatives are on-mission. For example, Adventist Health Lodi Memorial’s mission statement is “Living God’s love by inspiring health, wholeness, and hope.” One program they highlighted was a shoe drive for families in need, which while intuitively provides hope to these families, the text itself does very little to inspire hope, describing the event with plain details such as “The party included food and music.” In contrast, another program emphasized “dancing for health nutrition” and “wellness exercises,” aligning more closely with the mission statement and thus was predicted to have less deviation. Yet, a human reader could argue that both programs support the hospital’s mission statement equally. In cases like these, the model struggled as it would need to learn the intention behind ideas.

6 Conclusion

This study's primary contribution is the development of a novel measure for hospital mission deviation: lexical dissimilarity between annual reports and mission statements. Though there has been significant theorizing and qualitative research on mission deviation, there has been a relative dearth of formal measurement. This is a challenging measurement problem because there is no objective "ground truth" as to whether an organization's strategic priorities are consistent with its mission since a mission is inherently subjective to an organization's self-definition. One distinct strength of our metric is that it relies on an organization's activities *in its own words*. Rather than relying on external sources of information or the judgments of raters, we capture hospitals' descriptions of their own annual activities, and whether these actions are consistent with their own missions. We believe that this is the most robust and objective measure of mission deviation proposed thus far.

One of the most noteworthy findings of the current study is the positive association between OR and mission deviation, which suggests that revenue diversification may be harmful in maintaining a hospital's commitment to its mission. A central concern of Agency Theory is the alignment of the personal motivations of principles and agents in order to prevent agent opportunism or malfeasance; in other words, hospital leaders may make decisions that reflect their personal interests rather than the interests of the hospital's stakeholders. Therefore, according to agency theoreticians, revenue diversification can be viewed as the opportunistic pursuit of self-interest by poorly-monitored managers (Mark A. Fox, 1994). Opportunism and malfeasance are not isolated to the for-profit sector: for instance, LeClair (2019) defines the phenomenon of "soft corruption," or unethical (but not prosecutable) misappropriation of donated funds for personal gain. Though the results of the current study are too abstract to conclude that soft corruption is occurring in any of these hospitals, the association between mission deviation and revenue diversification does suggest that wasteful spending or inefficient use of capital may be occurring in some hospitals.

The scope of this study was limited by the availability of data. Future research in this area should aim to develop larger datasets that include multiple US states since this would provide additional statistical power and enable results to generalize to broader contexts. Additionally, human-labeled data in which trained raters manually scored the mission deviation by reading annual reports and mission statements would provide a more robust ground truth for the analyses in this study. Finally, though a time-series analysis was beyond the scope of the current study, as organizations publish annual reports each year, historical data could be used to determine mission deviation trajectories and identify factors that mitigate or exacerbate mission deviation across time.

Recent crises have highlighted the need for sound empirical guidance for practitioners in the nonprofit sector. Though the landscape of nonprofit financing is perennially unstable, the COVID-19 pandemic has been a period of pronounced volatility characterized by dramatic shifts in charitable giving and grant availability (Artemis Strategy Group). In such an unpredictable fiscal environment, nonprofit hospitals require sound empirical guidance in order to prevent mission deviation and its consequences, such as wasteful spending or administrative bloat. Additionally, policymakers in recent years have increasingly voiced concerns that nonprofits are not providing sufficient community benefit to justify their tax-exempt status (Bradley Herring, 2018). We hope that the results of the present study can provide additional insights to practitioners and policymakers in navigating these complex policy decisions.

References

- Christoph Gebele Dieter K. Tscheulin Ann-Kathrin Seemann, Florian Drevs. 2015. Are religiously affiliated hospitals more than just nonprofits? a study on stereotypical patient perceptions and preferences. *Journal of Religion and Health*, 54:1027–1039.
- Artemis Strategy Group. *COVID-19 and Philanthropy: How Giving Behaviors Shifted During the Pandemic*. Cincinnati, OH: Fidelity Charitable.
- Charles A. Barragato. 2002. Linking for-profit and nonprofit executive compensation: Salary composition and incentive structures in the u.s. hospital industry. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 13(3):301–311.
- Jessica L. Berrett. 2021. Linking overhead expenses and nonprofit effectiveness: Evidence from habitat for humanity. *Nonprofit Management and Leadership*, 32(4):509–530.
- Hossein Zare Gerard Anderson Bradley Herring, Darrell Gaskin. 2018. Comparing the value of nonprofit hospitals' tax exemption to their community benefits. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 55:1–11.
- Dilek Cetindamar and Banu Ozkazanc-Pan. 2017. Assessing mission drift at venture capital impact investors. *Business Ethics: A European Review*, 26(3):257–270.
- Irene Bloemraad Cheng Ren. 2022. New methods and the study of vulnerable groups: Using machine learning to identify immigrant-oriented nonprofit organizations. *Socius: Sociological Research for a Dynamic World*, 8:1–14.
- Christopher Cornforth. 2014. Understanding and combating mission drift in social enterprises. *Social Enterprise Journal*, 10(1):3–20.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding.
- Kirsten A. Grønberg. 2001. The u.s. nonprofit human service sector: A creeping revolution. *Nonprofit and Voluntary Sector Quarterly*, 30(2):276–297.
- Linda Lampkin. 2010. What nonprofits, grantmakers, and donors need to know. *Automatic Revocation of Nonprofits' Tax-Exempt Status*, pages 1–8.
- Mark Steven LeClair. 2019. Malfeasance in the charitable sector: Determinants of 'soft' corruption at nonprofit organizations. *Public Integrity*, 1:54–69.
- Arpan Mandal, Raktim Chaki, Sarbajit Saha, Kripabandhu Ghosh, Arindam Pal, and Saptarshi Ghosh. 2017. Measuring similarity among legal court case documents. In *Proceedings of the 10th Annual ACM India Compute Conference*, Compute '17, page 1–9, New York, NY, USA. Association for Computing Machinery.
- Robert T. Hamilton Mark A. Fox. 1994. Ownership and diversification: Agency theory or stewardship theory. *Journal of Management Studies*, 31(1):69–81.
- Sizakele Nkabinde and Anastacia Mamabolo. 2022. The south african social enterprises' strategies to guard against mission drift when faced with tensions from the funders. *Cogent Business & Management*, 9(1).
- Tommaso Ramus and Antonino Vaccaro. 2014. Stakeholders matter: How social enterprises address mission drift. *Journal of Business Ethics*, 143(2):307–322.
- Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks.

7 Appendix

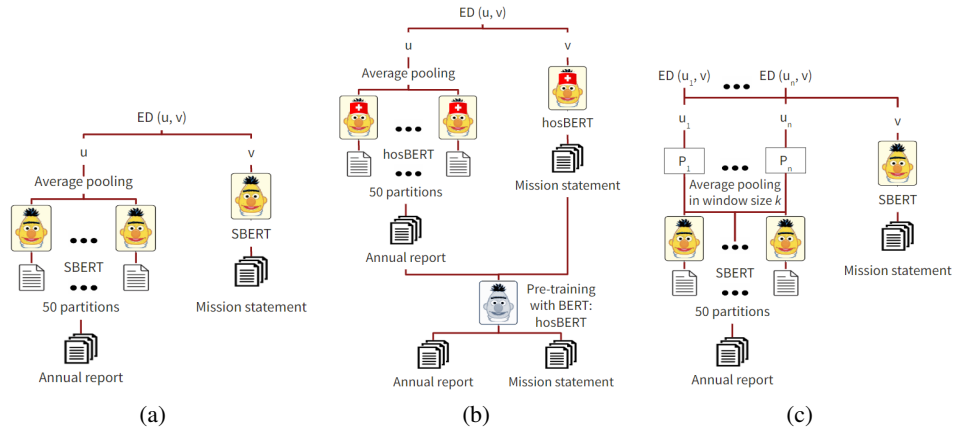


Figure 2: Embedding strategies: full document comparison with (a) SBERT (b) hosBERT and (c) progressive document comparison with SBERT