

Untangling The Program Name Versus The Curriculum: An Investigation of Titles and Curriculum Content

R. Wes Crues
University of Illinois
Dept. of Educational Psychology
1310 South Sixth Street
Champaign, Illinois
crues2@illinois.edu

ABSTRACT

This investigation focuses on the relationship between skills taught during business programs and whether the skills taught relate to the title of the program, as deemed by subject-matter experts. We hone-in on formal degree and non-degree programs in small business education, entrepreneurship education, or a blend of these two to determine if the *name* of the program is related to the *skills taught* in said program. We use a collection of excerpts from college catalogs, which are all descriptions of the formal academic programs. We then use k -means clustering to group program descriptions into interpretable clusters. We discuss the findings from the cluster analysis.

Keywords

text mining, clustering, higher education, business education

1. INTRODUCTION

Major academic disciplines are typically collections of finer-grained specialties; for example, a computer science department might consist of experts in human-computer interaction, artificial intelligence, algorithm design, among others. Colleges likely have departments with similar names, but we want to understand if similarly named degree programs at different universities equip students with similar skills. To discern whether or not this task is tractable, we used a collection of program descriptions from college catalogs about programs claiming to teach students entrepreneurship, small business, or a blend between these two curriculum areas. These definitions are used throughout:

- A *program description* is at least one, but often composes a few paragraphs, which delineates skills taught in programs, and might provide some learning goals and a listing of courses;
- *Entrepreneurship* is defined as “trying to identify opportunities and putting useful ideas into practice” [1]

Table 1: Distribution of Program Descriptions

Program Label	N	Degree/Non-Degree
Entrepreneurship	444	247/197
Small Business & E-ship	82	42/40
Small Business	79	20/59
Special Focus	92	34/58

(p. 6);

- and, *small business management* is “the ongoing process of owning and operating an established business” [3] (p. 28).

Our study explores whether we can use text clustering to identify a clear distinction between these two areas of business education, determine if there are differences between two-year and four-year programs, and whether there are differences between degree and non-degree programs.

2. METHOD

A research team manually assembled a collection of 697 program descriptions from college catalogs for institutions located in the United States. Research assistants went to college websites and manually extracted text from published college catalogs online. The initial list of programs was derived from the 2013 Integrated Postsecondary Education Data System (IPEDS) maintained by the United States Department of Education. After filtering institutions which did not have any business programs, a random sample of programs arrived at the collection used.

Program descriptions spanned programs focusing in entrepreneurship, small business management, or a blend of the two. Additional program descriptions were collected which were considered special focus programs; these were programs which teach a specific skill set on operating a business (examples include funeral home management to hair weaving and braiding entrepreneur). We also considered formal degree (e.g., associates and bachelor degrees) or non-degree programs (e.g., certificates or specializations), and whether the home institution is public or private, for-profit or not-for-profit, and whether the institution is a 2-year, 4-year, or 4-year and beyond institution [5]. Table 1 presents the distribution of program labels and whether the program is a degree or non-degree program.

2.1 Preprocessing Program Descriptions

Program descriptions were transformed into raw text format, tokenized into unigrams, except for a few words. A few bigrams and trigrams were specified using knowledge from a domain-expert, for example, business plan(s), social entrepreneurship, home based business, and venture capital. Punctuation, numbers, and top words were removed using the pre-defined English stop word list in the “tm” package in R [2]. We used stemmed words by using the Porter stemming algorithm [6]. We used binary indicators to determine whether a term was present in each program description when constructing the document-term matrix [4].

2.2 Corpus Statistics

Our initial document-term matrix contained 7799 unique terms with a sparsity of 99%. We removed very frequent terms deemed to have no substantive value by a domain expert. Due to the nature of the corpus (i.e., program descriptions), words such as catalog, college, semester, requirements, and introduction, among others, were excluded. Eventually, we used the “removeSparseTerms” function in the “tm” package in R [2], which resulted in a document-term matrix with 16 unique terms, however, still 70% sparse.

2.3 Program Description Clustering

We utilized k -means because this clustering technique was favored in prior studies [7]. We experimented with various numbers of centroids, and after discussions with domain experts, we determined $k = 10$ was an optimal solution. The domain expert believed this solution provided an interpretable and reasonable grouping of programs. Specifically, the distribution of whether the program was an entrepreneurship, small business, a blend of these, or a special focus program, coupled with their expectations of distribution of formal degree programs versus non-degree programs. More than ten centroids resulted in clusters containing less than five documents, while less than ten resulted in a solution which did not provide what domain experts believed to be the most interpretable.

3. RESULTS

Five of the clusters exhibited a focus on teaching entrepreneurship in the context of having an idea, creating a start-up, with the intention of scaling the business into a large enterprise. Within these clusters, two clusters had words indicating programs might teach entrepreneurship to equip students to solve global problems and health concerns. Words indicating entrepreneurship might be taught to professionals in fields besides business (i.e., law and engineering) appeared in one cluster. One cluster appeared to teach general business skills, without a clear focus on entrepreneurship or small business. Another cluster contained special focus programs, which seek to prepare students for a specialized, technical career, such as a travel agent or carpenter. Two clusters contained small business programs, where one focused on keenly on running ones’ own business, while the other included this while teaching students to innovative. One cluster contained very detailed program descriptions from one institution.

4. DISCUSSION & CONCLUSIONS

We found the definition of entrepreneurship which pertains to creating and expanding new enterprise appeared to be

almost exclusively in four-year colleges, especially research universities. In contrast, small business management and operating a small business were taught almost exclusively at two-year colleges. A few of the two-year colleges also had many specialized programs in applied fields, such as the cosmetology; these types of programs were nearly exclusive to two-year colleges. Another element of entrepreneurship is creativity and innovation. These skills, specifically innovation, seemed to be taught primarily in the four-year sector. The programs that considered themselves a blend tend to focus more on small businesses than entrepreneurship. We found innovation and these skills to be taught more in degree. On the other hand, skills related to managing a small business were in non-degree programs.

From our findings about entrepreneurship and small business education, we generally found labels of programs match the skills one would expect to learn given the name of the program. However, one cluster in our analyses did not indicate skills in the targeted areas were being specifically taught. A limitation of our study is program descriptions vary in length and detail, which might be problematic for clustering. Our further work plans to consider whether skills taught have changed over time; for example, are skills being taught today the same skills taught a decade ago?

5. ACKNOWLEDGMENTS

The author would like to acknowledge two domain experts, Dr. Cindy Kehoe and the late Dr. Paul Magelli for their expertise in entrepreneurship education. Their advice about contextual meaning of results was invaluable in interpreting these analyses. The author would also like to acknowledge the Ewing Marion Kauffman foundation, which funded this work through a grant to inventory and interpret entrepreneurship education in higher education in the United States.

6. REFERENCES

- [1] B. R. Barringer and R. D. Ireland. *Entrepreneurship: Successfully launching new ventures*. Pearson, Upper Saddle River, New Jersey, fourth edition, 2012.
- [2] I. Feinerer, K. Hornik, and D. Meyer. Text mining infrastructure in r. *Journal of Statistical Software*, 25(5):1–54, 2008.
- [3] T. S. Hatten. *Small business management: Entrepreneurship and beyond*. Houghton Mifflin Company, Boston, Massachusetts, fourth edition, 2009.
- [4] P. Howland and H. Park. Cluster-preserving dimension reduction methods for efficient classification of text data. In M. W. Berry, editor, *Survey of Text Mining*, pages 3–24. Springer Science+Business Media, 2004.
- [5] National Center for Education Statistics. *IPEDS Glossary*, 2017.
- [6] M. F. Porter. An algorithm for suffix stripping. *Program: Electronic Library and Information Systems*, 40(3):211–218, 2006.
- [7] M. Steinbach, G. Karypis, V. Kumar, et al. A comparison of document clustering techniques. In *KDD Workshop on Text Mining*, volume 400, pages 525–526. Boston, 2000.