Automated extraction of learning goals and objections from syllabi using LDA and neural nets.

Poster · March 2018

7 authors, including:

Rinat Levy-Cohen
Fordham University
14 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Automated Extraction of Learning Goals and Objectives from Syllabi using LDA and Neural Nets

Charles Lang
Columbia University
cl3584@tc.columbia.edu

Rinat Levy-Cohen, Charlotte Woo, Benjamin Roberts, Samantha Pepe, Roshan Verma, & Yan Xu

ABSTRACT: Effective application of learning analytics to the classroom requires the generation of reliable standards on which to base metrics (Ferguson, 2012). An effective, but highly labor intensive way to create these standards is for instructors to define competencies, learning goals and objectives for their course(s). Although this may be the ideal solution, it may be unrealistic to scale the approach, especially for courses that have already been established for many years. Therefore, there is a need for technological tools that can streamline this process. The following research describes a process for leveraging syllabi that do not have learning standards defined to automatically extract learning goals and objectives. The process utilizes Latent Dirichlet Allocation (LDA) topic modelling coupled with neural nets to generate sentences describing learning goals and objectives directly from a syllabi for review by an instructor or course designer.

Keywords: competency based education, learning goals, Latent Dirichlet Allocation (LDA), topic modelling, Bloom's taxonomy, text imputation, neural net.

1 INTRODUCTION

Competency based learning is a model characterized by explicitly defined knowledge and skills, expediency in results and close alignment to industry demands (Voorhees, 2001). Tightly-defined learning goals and objectives become pertinent in ensuring that competency based learning achieves the intended outcome. Good learning goals and objectives are difficult and time consuming to generate (Piskurich, 2015). Any effort to automate the process of creating goals and objectives without compromising their quality will be valuable to educators and designers and crucial for feeding into analytic systems. This work builds on Ramesh, Sasikumar, & Iyer (2016) who developed an automated system to integrate course content and levels of thinking (as defined by the syllabus) into a learning objective annotated ontology (LAO) and connects to other efforts to aid instructional design such as Avila et al. (2017) and Greer et al. (2015).

2 METHODS

Seventy syllabi were collected using convenience sampling. Variables from the headings were extracted as learning goals, and those from the body as learning objectives using the PDFBOX and Cloud Java libraries. Text was processed and Latent Dirichlet Allocation (LDA) was performed to extract topics from this corpus using the tm, Snowball, NLP and tidytext R packages. Topics were

then defined as the most common word per topic, and topic names were manually matched to objective verbs from Bloom's hierarchy of learning objectives. (Hsiao & Awasthi, 2015). A three-layer neural net was then trained to predict which learning objectives matched with the generated topics using the neural net R package. The trained neural net was tested on a sample of three syllabi.

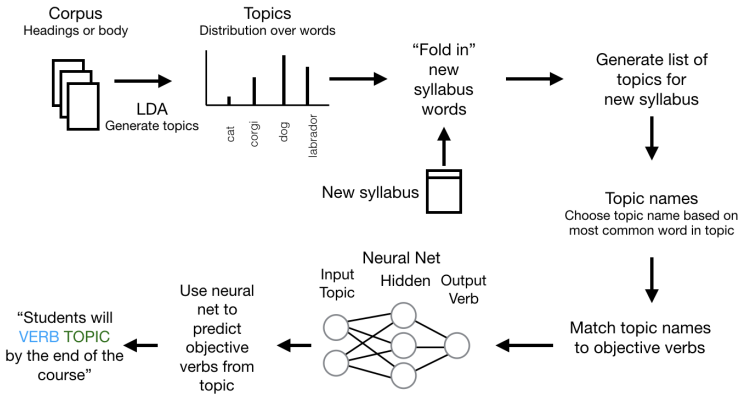

Figure 1: LDA and neural net process flow

3 RESULTS & DISCUSSION

In the future, we hope to judge how well the system works by collecting data on the needs of the syllabus author. We plan to create a system that we can make available to instructors and capture this information. We believe that there is utility for the learning analytics community in using such systems to aid the creation of useful standards and meaningful metrics in syllabi learning goals.

4 REFERENCES

Avila, C., Baldiris, S., Fabregat, R., & Graf, S. (2017, March). ATCE: an analytics tool to trace the creation and evaluation of inclusive and accessible open educational resources. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 183-187).

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5–6), 304–317.

Hsiao, I. H., & Awasthi, P. (2015). Topic facet modeling: semantic visual analytics for online discussion forums. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge.* (pp. 231-235).

Piskurich, G. M. (2015). *Rapid instructional design: Learning ID fast and right*. New York: Wiley & Sons.

Ramesh, R., Sasikumar, M., & Iyer, S. (2016, July). Integrating the learning objectives and syllabus into a domain ontology for data structures course. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 266-271).

Voorhees, R. A. (2001). Competency-Based learning models: A necessary future. *New directions for Institutional Research*, 2001(110), 5-13.

Greer, J., Molinaro, M., Ochoa, X., & McKay, T. (2016, April). Learning analytics for curriculum and program quality improvement (2016). In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 494-495).