Extracting and evaluating educational concept dependencies implicit in Large Language Models

Dominik Glandorf
Matrikelnummer 6007407
dominik.glandorf@student.uni-tuebingen.de

Anastasiia Alekseeva
Matrikelnummer 5994775
anastasiia.alekseeva@student.uni-tuebingen.de

GitHub repository: https://github.com/mlcolab/learning-dependencies

Abstract

1 Introduction

Large Language Models (LLMs) are trained on immense corpora of text and have proven to build on included factual information when performing well in downstream tasks such as question answering. Accessing this knowledge represented by billions of parameters and the network's architecture has given rise to the research field of Knowledge Extraction from LLMs (Cohen, 2023). Here, we focus on educational knowledge to create new perspectives for a subfield of knowledge engineering that has rather stagnated in the previous years.

Concept dependencies are complex but highly relevant educational knowledge. Effective and efficient instruction does not only incorporate what to teach but also how to teach. Guidelines for *instructional sequencing* emphasize the order of instruction. More precisely, prerequisites of educational content should be either known to the student or taught first [Morrison et al., 2019]. Within educational content, Merill (1983) differentiated facts, concepts, principles, rules, procedures, interpersonal skills, attitudes, and their sole recall from their application. To simplify, we will focus on concepts and their relations in terms of understanding. If one concept is a prerequisite of another, we call this relation a concept dependency. For example, to understand the concept of a derivative, having knowledge about the concept of a function will facilitate or even enable learning. When the dependencies are thought of as directed edges between nodes that represent concepts, a concept dependency graph emerges which is a special type of a knowledge graph [Wang et al., 2016]. This graph is also called *concept map* in the field of Learning Sciences. The graph can be used to advance curriculum planning (Yang, 2015), especially for new topics that might not be covered in textbooks or automated assessment (Wang, 2015).

In this work, we tackled the question how to extract this particular knowledge graph and how to evaluate its quality.

First, there was no classical benchmark available to evaluate the accuracy of extracted information. Therefore, we propose a set of methods to use existing unstructured knowledge sources, namely Wikipedia and textbooks, to create baselines for evaluation. Due to the heuristic character of these methods, we conducted a manual assessment to test their suitability for our purpose. The resulting dataset can be used as a human baseline for further research.

Second, there is no established manner to extract the desired knowledge from the LLM. The emerging field of prompt engineering is currently a vast collection of commands to query LLMs. We propose a method called output refeeding that sequentially queries the language generation model and transforms its answers into a knowledge graph.

1.1 Related work

Prerequisites can be inferred from learner behavior by testing their performance after being presented different instructional sequences (Pavlik et al., 2008, Vuong et al., 2011). However, this has the disadvantage of disengaging users with too difficult concepts before teaching easier or necessary ones. Experts usually dispose of the required knowledge about concepts to create concept maps. The high cost of expert knowledge motivates the automated extraction of concept dependencies from appropriate sources.

Talukdar and Cohen [2012] defined the prerequisite relation in terms of the consumption of information about concepts. Vuong (2011) if a better graduation rate given prerequisite knowledge is fulfilled. Concepts are often equated with Wikipedia articles (Talukdar and Cohen, 2012; Wang, 2015).

2 Method

In this section, we will first detail our research design and then the characteristics of our information sources as well as the methods that we used to produce the knowledge graph.

2.1 Research design

2.2 Baseline extraction

2.2.1 Wikipedia

2.2.2 Textbooks

For preprocessing the textbooks we used Wikifier [Brank et al., 2017].

2.3 LLM extraction

2.4 Manual inspection via Dashboard

2.5 Manual baseline

2.6 Convergence statistics

3 Results

4 Discussion

References

Gary R Morrison, Steven J Ross, Jennifer R Morrison, and Howard K Kalman. *Designing effective instruction*. John Wiley & Sons, 2019.

Shuting Wang, Alexander Ororbia II, Zhaohui Wu, Kyle Williams, Chen Liang, Bart Pursel, and C Lee Giles. Using prerequisites to extract concept maps from textbooks. In *Proceedings of the 25th acm international on conference on information and knowledge management*, pages 317–326, 2016.

Partha Talukdar and William Cohen. Crowdsourced comprehension: predicting prerequisite structure in wikipedia. In *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pages 307–315, 2012.

Janez Brank, Gregor Leban, and Marko Grobelnik. Annotating documents with relevant Wikipedia concepts. *Proceedings of SiKDD*, 472, 2017.