

Using NLP to Measure Relations Between Different Philosophers and Philosophical Schools of Thought

Bobby Becker
Tulane University
rbecker3@tulane.edu

Problem Overview

In philosophy, there are many different threads of thinkers who have built their ideas with the works of earlier thinkers; the relationship between different ideas and different thinkers is well understood in a qualitative sense, but with modern NLP tools, we can validate those associations in a more objective way. We aim to extract key themes, topics, and philosophical concepts from these texts to gain insights into the philosophical ideas presented by different thinkers, and on how different philosophical thinkers influence each other. Specifically, we will measure the influences and overlap of Ancient Greek philosophers—Plato, Aristotle, and the pre-Socratics—and European philosophers, leading up to and resulting from Kant. These works are of particular importance since they have played a huge role in influencing the development of the scientific, ethical, and political frameworks of the Western world.

Data Collected

The data that we will collect for this project includes the corpora of key philosophers (translated to English), including Plato, Aristotle, Kant, Leibniz, Descartes, and others. This data is collected through Project Gutenberg, which provides the full texts of many of those philosophers as part of the public domain. In addition, we will also be collecting the Wikipedia pages of those

philosophers. The data will be preprocessed to eliminate filler words and to tokenize only meaningful words mentioned in their philosophical works and their Wikipedia pages.

Algorithms and Methodology

We will be creating a Latent Dirichlet Allocation (LDA) model on the various texts of major philosophers—creating individual models for individual texts and models for the entire corpus of various philosophers. The LDA model is an unsupervised machine-learning model that can generate the main topics from each set of texts. With topics generated, we can then compare what are the main topics covered by each philosopher, giving us some baseline data to start making observations. In addition, we will also be employing a Word2Vec bag-of-words models on the entirety of the corpora collected. With this data, we will create vector representations of each philosopher and can measure the similarity between different philosophers and different philosophical schools. Finally, we will be employing a Word2Vec model on the Wikipedia pages collected to have additional text data which we can use to measure the similarity between different philosophers.

Preliminary Results

Up to this point, we have successfully preprocessed and created LDA & Word2Vec models for the main three works of Kant: *The Critique of Pure Reason*, *The Critique of Practical Reason*, and *The Critique of Judgement*. With the LDA model, were able to measure the main topics of these three works in a very accurate way. With the Word2Vec model, we were able to create somewhat accurate associations between different ideas. We also successfully fetched, preprocessed, and created a Word2Vec model for 500 Wikipedia pages related to different philosophers; our Word2Vec model successfully shows which ideas are associated with each

other and which philosophers are related to each other. However, we likely would want to experiment with training on a larger Word2Vec model trained on a larger portion of Wikipedia to get more accurate results.

Related Works

[Domain-Specific Evaluation of Word Embeddings for Philosophical Text using Direct Intrinsic](#)

[Evaluation](#): This study evaluates the effectiveness of word embeddings trained on the texts of a single philosopher to capture domain-specific meanings of philosophical terms. Traditional language models trained on vastly more texts often do not accurately represent the nuanced language found in humanities texts, including philosophy; the researchers here compare six models trained on specific philosophers to see which one best aligns with expert judgments in tasks related to synonym detection and coherence. While this study is a great example of how Word Embeddings can be studied on the texts of individual philosophers, it is aimed at creating a generative model, and the authors did not use their research to do any comparative analysis between the philosophers studied.

[LDA Topic Modeling: Contexts for the History & Philosophy of Science](#): This paper explores the use of Latent Dirichlet Allocation (LDA) topic modeling in the fields of history and philosophy of science (HPS), using the study of Charles Darwin's reading and writing habits as a primary example. It addresses common misconceptions about topic models, arguing that they offer more than just improved search capabilities by revealing contextual insights into how figures like Darwin engaged with their reading materials, thereby opening new avenues for investigating intellectual influence and creativity. My project can be seen as a particular application of this research on the philosophers that I have focused in on studying.

[Measuring Philosophy in the First Thousand Years of Greek Literature](#): In this study, the author demonstrates how LDA topic modeling can be used to automatically identify philosophical passages within a large corpus of the first thousand years of Greek literature, sourced from the Open Greek and Latin group and the Perseus Digital Library. By combining qualitative analysis, topic modeling, and expertise in Classics, the study devises three numerical scores to evaluate texts on "good and virtue," "scientific inquiry," and their overall "philosophicalness." My work is a way to apply similar methods to this to both Greek and European philosophers.

[Eight journals over eight decades: a computational topic-modeling approach to contemporary philosophy of science](#): This study applies computational text-mining and topic-modeling algorithms to the full-text content of nearly 16,000 articles from eight major philosophy of science journals, spanning from the 1930s to 2017, to trace the evolution of the discipline's research themes. Through this analysis, 25 research themes and 8 thematic clusters were identified, illustrating the dynamic changes in the philosophy of science's research agenda over eight decades and how each journal uniquely contributed to these thematic developments. My research uses similar algorithms, but in an attempt to measure the development of specific pivotal figures who have contributed to the development of science and philosophy, not of journals.

Mapping the Modern History of the Philosophy of Religion with Machine Learning: This study uses the LDA model to map the development of the philosophy of religion, proposing it as a new way to revitalize scholarship in the field. Like the study above, however, this research studies the development of the field using data from journals, not from the source texts.

Timeline:

April 1st: 20 corpora loaded in and preprocessed. 1,000 Wikipedia pages loaded in and preprocessed.

April 8th: 100 LDA models tested. 10 significant data points collected.

April 15th: 100 Word2Vec models tested. 20 significant data points collected.

April 22nd: Models tweaked, finalized, and visualized..

April 30th: Final report completed.

References:

Foster, J.C. (2020). Mapping the Modern History of the Philosophy of Religion with Machine Learning. *The Macksey Journal*. Retrieved from <https://mackseyjournal.scholasticahq.com>

Malaterre, C., Lareau, F., Pulizzotto, D., & St-Onge, J. (2021). Eight journals over eight decades: a computational topic-modeling approach to contemporary philosophy of science. *Synthese*. Springer.

Köntges, T. (2020). Measuring philosophy in the first thousand years of Greek literature. *Digital Classics Online*. Retrieved from journals.ub.uni-heidelberg.de

Chandra, R., & Ranjan, M. (2022). Artificial intelligence for topic modelling in Hindu philosophy: Mapping themes between the Upanishads and the Bhagavad Gita. *PLOS ONE*. Retrieved from journals.plos.org

Allen, C., & Murdock, J. (2020). LDA topic modeling: Contexts for the history & philosophy of science. Retrieved from philsci-archive.pitt.edu

