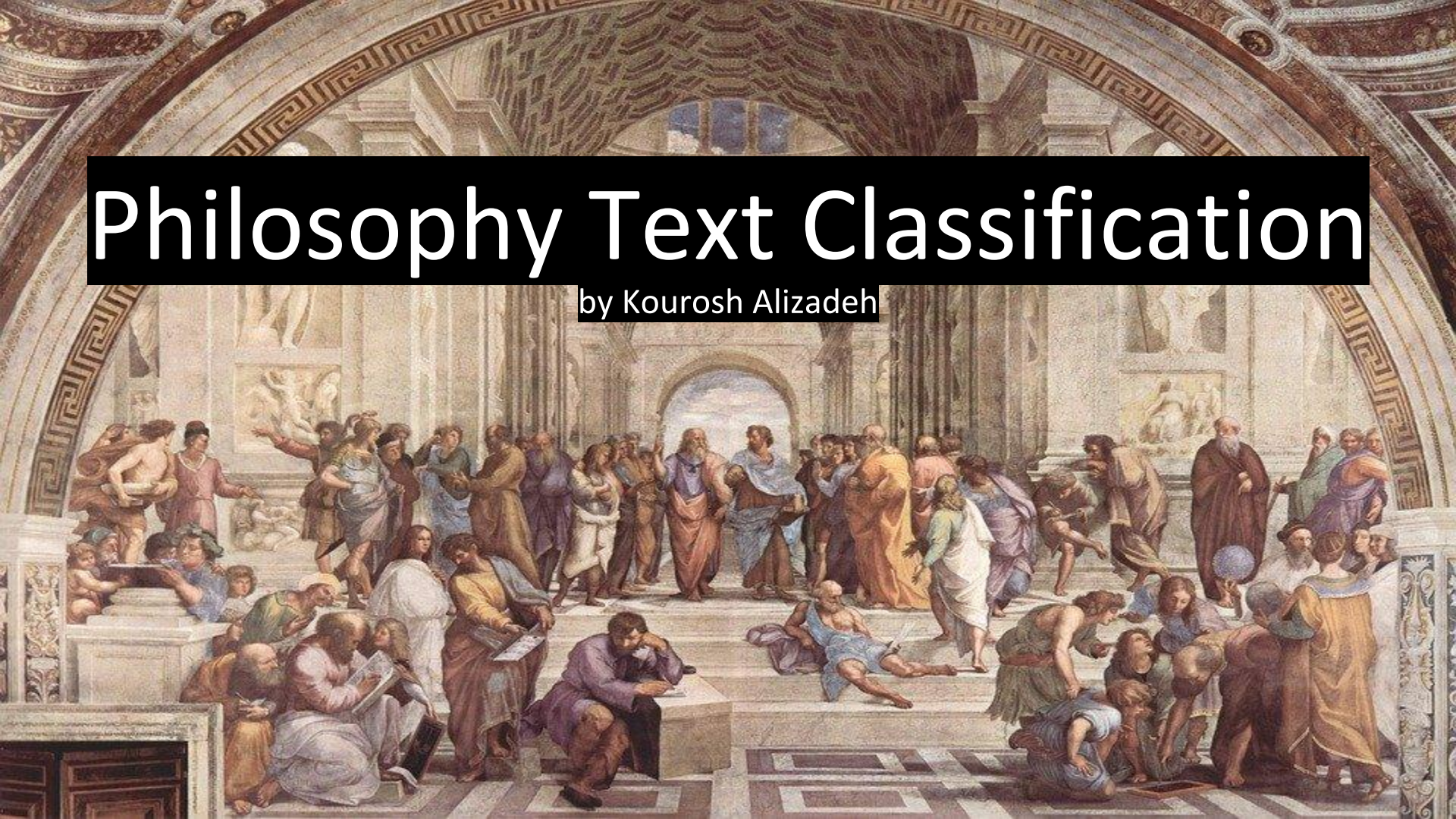


Philosophy Text Classification

by Kourosh Alizadeh



Why Classify Philosophical Texts?

- A philosophy represents a rationally developed worldview
- If we can classify texts based on their philosophical schools, we can classify sets of rationally organized perspectives on life
- This can be used in a purely academic way to isolate key features of a thinker's thought or notice differences hitherto unexplored
- If we expand this to the general populace, it means we can take user-generated text (tweets, FB posts) and identify a general worldview
 - this could be used in marketing, political campaigning, or sociological research

Word Clouds for Each School of Thought

Plato Word Cloud

think things man friend
men others see say
one take seem
people name
SOCRATES
law case kind tell
come soul know like time

Aristotle Word Cloud

reason motion
things
say belong animal possible body
part either form men
one come kind
good state
like others nature place cause

Rationalism Word Cloud

things one
truth sense power nothing cause
God idea
nature think good true reason
say motion certain order man
mind know body
soul object without men see love part

Empiricism Word Cloud

object men part
man reason sense
idea without
thought time
passion knowledge
another knowledge principle nature things
mind action nothing name
word relation matter though power

German Idealism Word Cloud

object determination
relation time one
existence idea without unity
reason principle consciousness universal thought
end self
understanding cognition content general
nothing nature form law
moment world means

Phenomenology Word Cloud

essence consciousness sense
things one experience
knowledge possible perception form always
time object means
present existence thought subject truth
world fact say beings meaning question
within science become work

Analytic Word Cloud

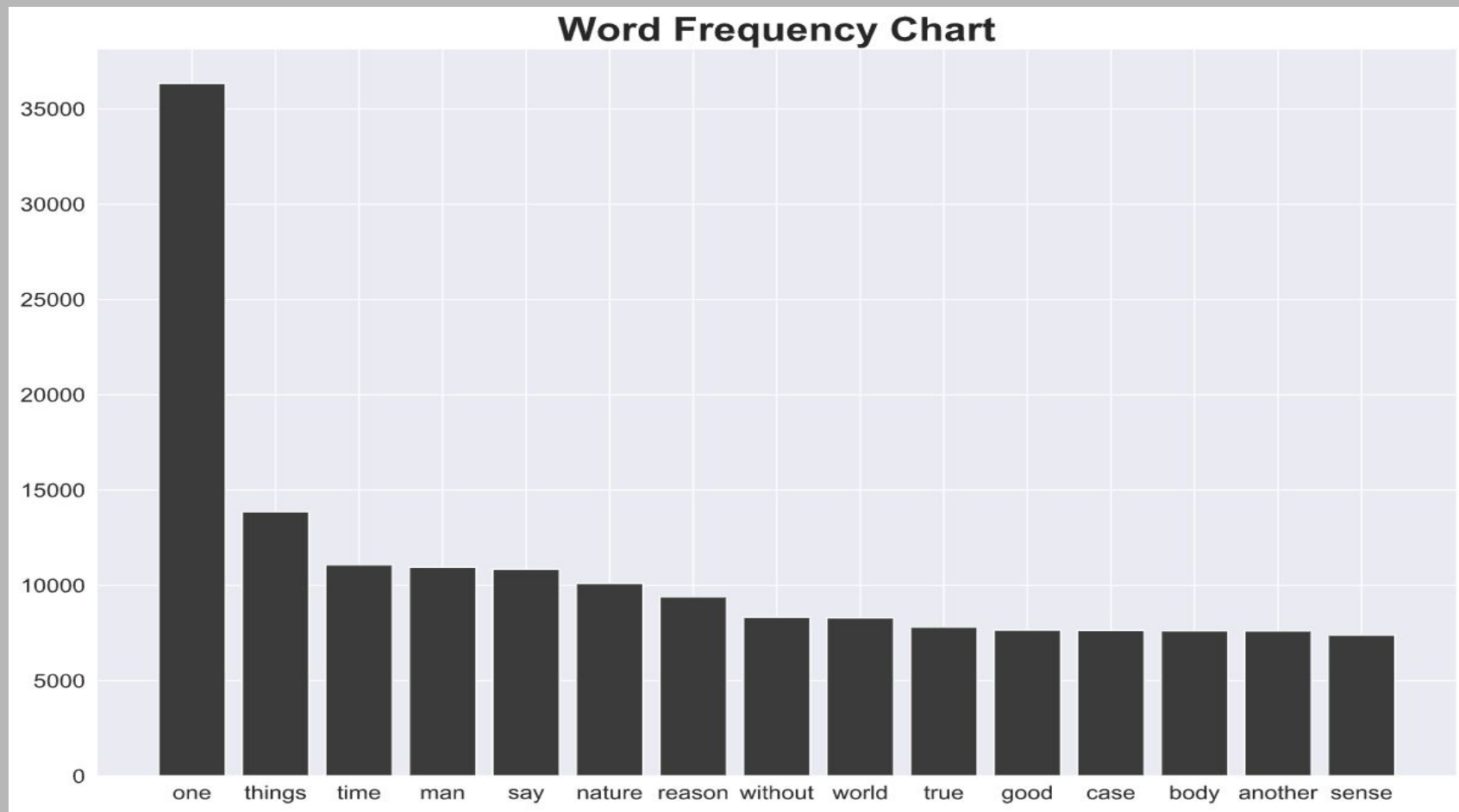
sentence world true
one see
use object term fact
sense time
proposition question say
like language seem
think theory word know
statement

Continental Word Cloud

point idea become
time one
possible object desire knowledge
relation order
question within reason
language madness
thought
man truth
nature world
representation difference word

page
left intentionally
this
blank

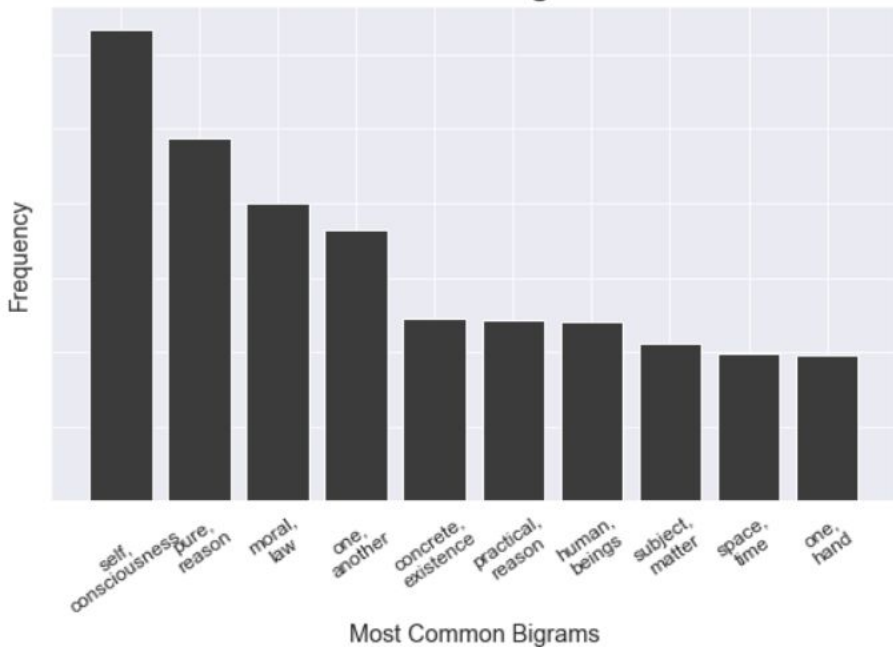
Overall Word Frequency



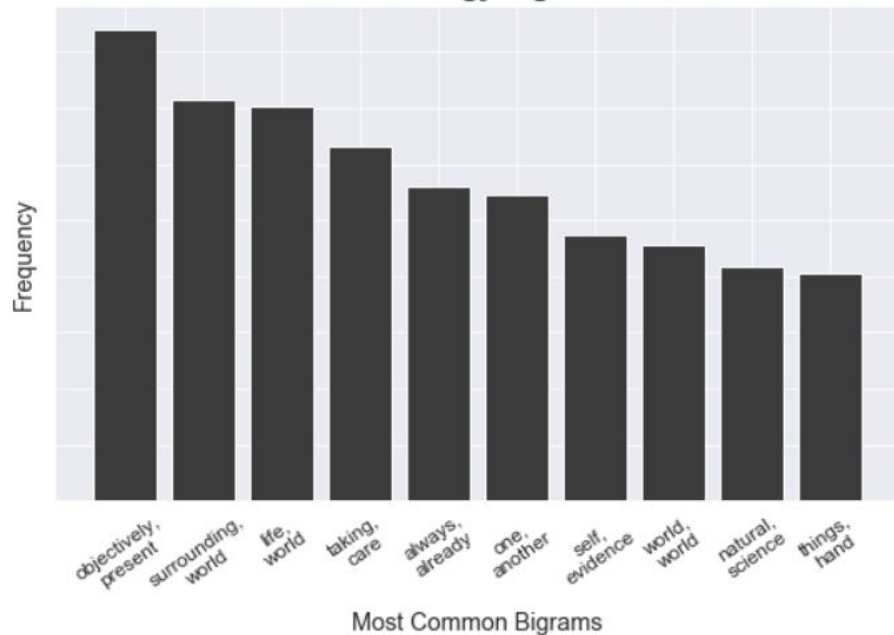
A Few Illuminating Bigrams

- self consciousness is a key concept in German Idealism
- the idea of objective presence and the phrase 'always-already' are common in Phenomenology

German Idealism Bigram Chart

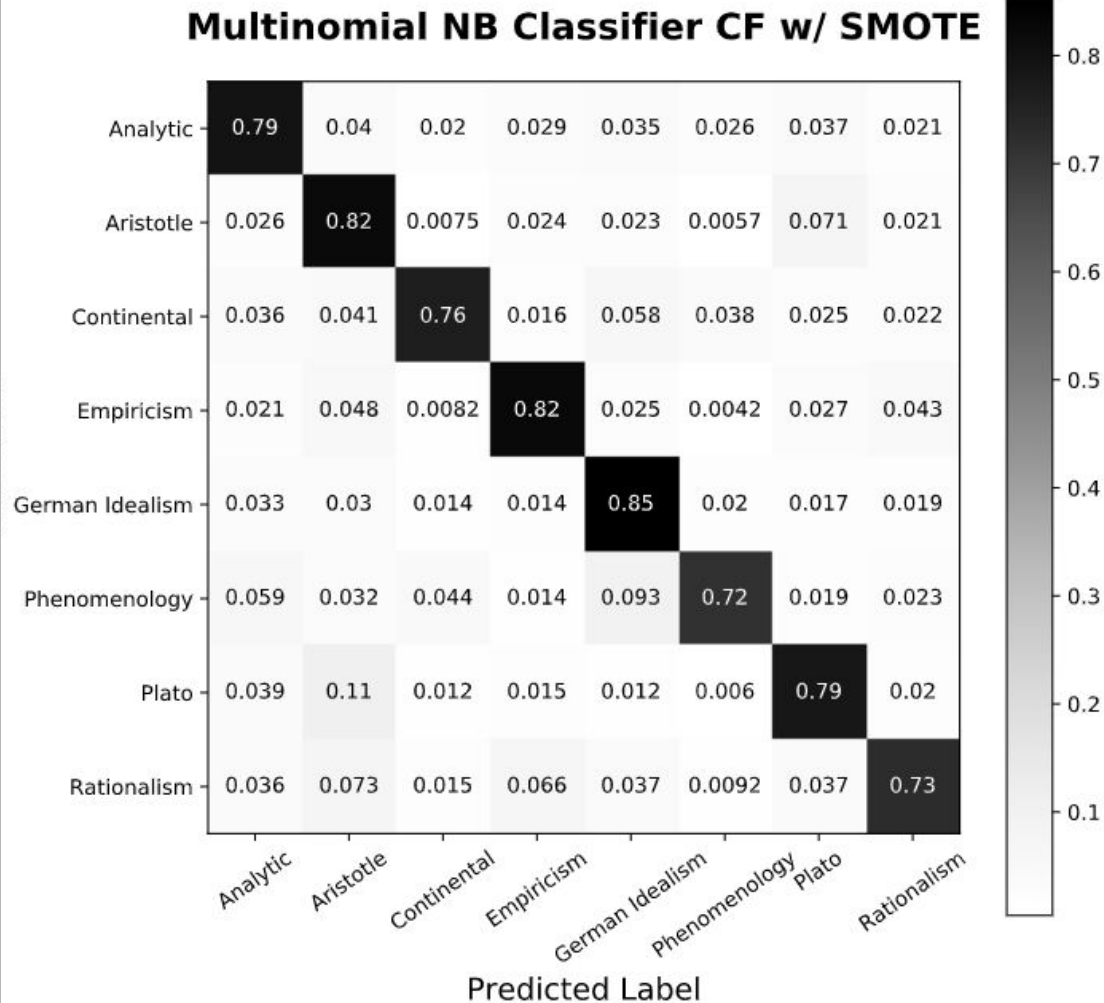


Phenomenology Bigram Chart



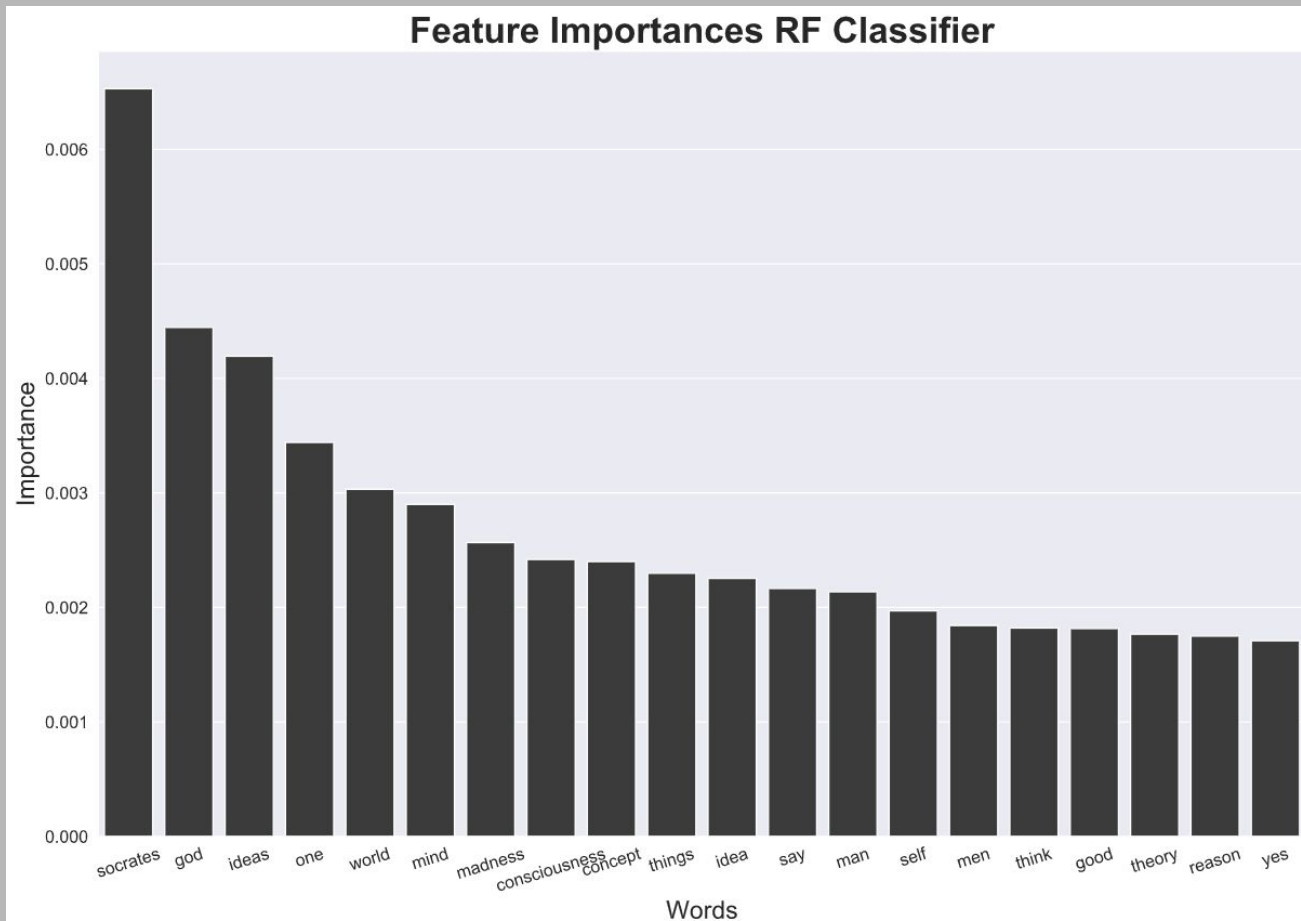
Bayesian Model

- Better performance than Random Forest
- 79% accuracy rate
- Aristotle, Empiricism, and German Idealism have strong recall rates.



Random Forest Model Feature Importances

- RF was less accurate but gave us feature importances.
- 'god' is a major topic in some schools but not others,
- 'socrates' is much more common in Plato than elsewhere.



Key Takeaways

- We were able to build a model with nearly 80% accuracy to classify philosophy texts by school
- Bigrams and word frequency charts effectively capture key terms and phrases of the schools
- Common words in all schools include 'one,' 'concept,' and 'time.'
- Key classificatory terms include 'socrates,' 'think,' 'god,' and 'madness.'
- Proof-of-concept operationalization in a notebook that classifies user text

Future Work

- Clean the data more thoroughly, especially for typos and word fusing that was common in the data (e.g., 'aconcept' for 'a concept').
- Develop more effective neural networks.
- Take the user-input notebook and set it up in a webapp.
- As an extreme goal, it would be nice to build a generative model that could create text in the style of a philosopher or paraphrase given text in the style or worldview of a given school.
 - this would enable us to generate text to fit a client's worldview after identifying it

Thank you