The Obsessions of the Wise Minds

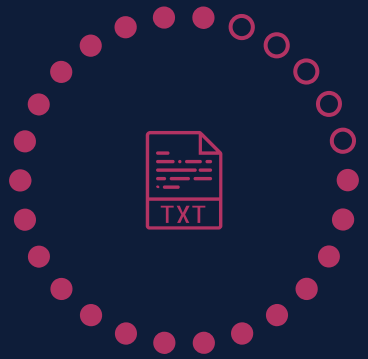
Topic Modeling of Over 2500 Years of Philosophical Texts

Milad Afrasiabi
November 2021

Overview & Problem

- Published texts fall into categories
- Wide variety of benefits
- Categories are hidden
- Are there naturally occurring topics and can they be discovered?

Proposal



Analyze Published Texts

Finding Naturally Occurring Topics

Using the Topics for Other Purposes

Pilot: Using Only Philosophical Texts

Beneficiaries

Libraries

Schools

Retailers

Online Services

Data

300,000 Sentences

Over 50 Original Texts

From Plato to Wittgenstein

From Kaggle

Methodology

Preprocessing

RE & String Libraries

Vectorization

TFIDF & Countvectorizer

Topic Modeling

NMF & LSA

Assessment

EDA & Domain Knowledge!

Raw Topics

```
Topic 0:
 FORM, SELF, CONSCIOUSNESS, THOUGHT, GENERAL, CONCEPT, IDEA, TRUTH, RELATION, PRESENT, MATTER, ORDER, PARTICULAR, SUB
JECT, CERTAIN

Topic 1:
 MAN, WOMAN, ANIMAL, WHITE, LOVE, PROPERTY, RIGHT, YOUNG, THINK, MAKES, HAVING, CERTAIN, BETTER, PERSON, INDIVIDUAL

Topic 2:
 NATURE, LAWS, NATURAL, ORDER, CAUSE, JUDGEMENT, FAR, FREEDOM, ART, CAUSES, ENDS, CONTRARY, PURPOSIVENESS, BODIES, UN
DERSTANDING

Topic 3:
 GOOD, BAD, EVIL, PLEASURE, LOVE, THINK, DESIRE, PLEASANT, RIGHT, VIRTUE, PAIN, MORAL, KIND, BETTER, SOUL

Topic 4:
 VALUE, LABOUR, CAPITAL, MONEY, QUANTITY, PRICE, PRODUCTION, PRODUCE, MEANS, SURPLUS, COMMODITIES, NECESSARY, WAGES,
COUNTRY, COMMODITY

Topic 5:
 REASON, PURE, PRACTICAL, UNDERSTANDING, EXPERIENCE, MADNESS, CRITIQUE, MORAL, PRIORI, SPECULATIVE, COGNITION, REGAR
D, THINK, FREEDOM, CAUSE

Topic 6:
 WAY, CERTAIN, RIGHT, THINK, KIND, THINKING, NEW, LET, SHALL, BEST, DASEIN, PARTS, WORK, POINT, ABLE

Topic 7:
 BODY, SOUL, MIND, PARTS, BODIES, PLACE, IDEA, CERTAIN, MOVEMENT, SPACE, REST, WATER, HEAT, ORGANS, AIR

Topic 8:
 OBJECT, CONCEPT, SUBJECT, EXPERIENCE, RELATION, CONSCIOUSNESS, REPRESENTATION, PERCEPTION, INTUITION, JUDGEMENT, COG
NITION, UNDERSTANDING, PLEASURE, DESIRE, MEANS

Topic 9:
 SENSE, COMMON, CERTAIN, WORD, PERCEPTION, DATA, SCIENCE, THINK, NATURAL, EXPERIENCE, TERM, KIND, WORDS, THEORY, MAKE
S

(10, 48809)
```

- NMF (10 Topics)
- CountVectorizer

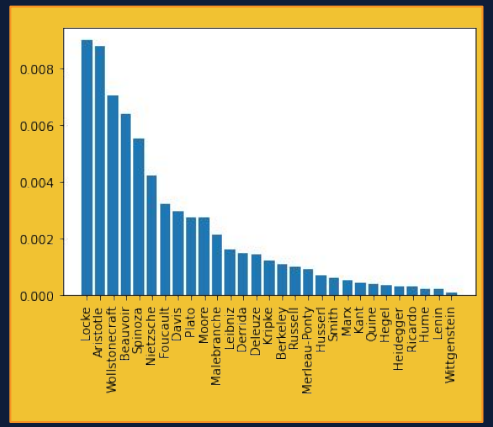
Topic Evaluation

- Author's share of the top 1000 excerpts in each topic (normalized)
- Top 5 authors for each topic (PCs)
- List of the most salient words in each topic
- Domain Knowledge!

author	Topic0	Topic1	Topic2	Topic3
Aristotle	0.004921	0.008412	0.004826	0.005092
Beauvoir	0.004016	0.012556	0.002008	0.001899
Berkeley	0.006885	0.002858	0.006096	0.001579
Davis	0.003380	0.004496	0.000924	0.000973
Deleuze	0.008916	0.002159	0.003814	0.001278
Derrida	0.007893	0.002491	0.001468	0.000619

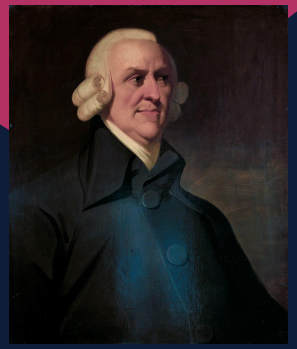
Topic Assessment: Topic 1

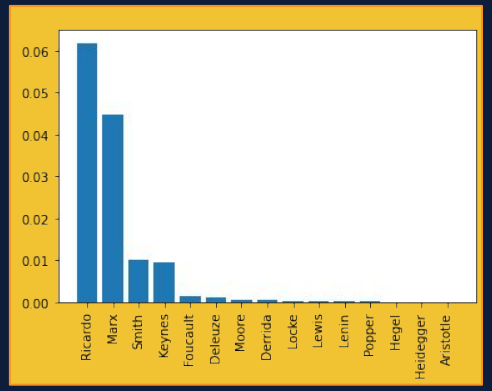
MAN, WOMAN, LOVE



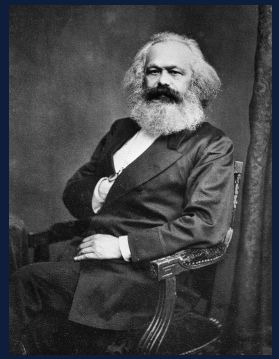
Feminism

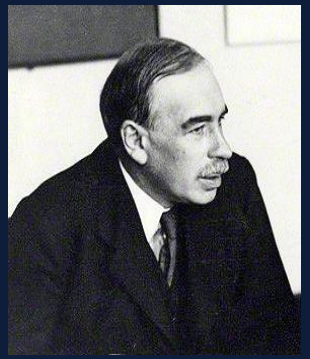
Topic Assessment: Topic 4





VALUE, LABOUR, CAPITAL, MONEY, SURPLUS, COMMODITY, WAGES

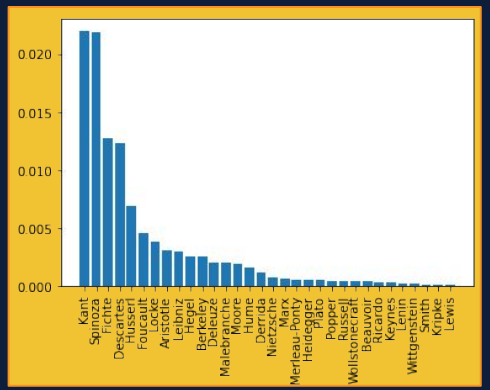


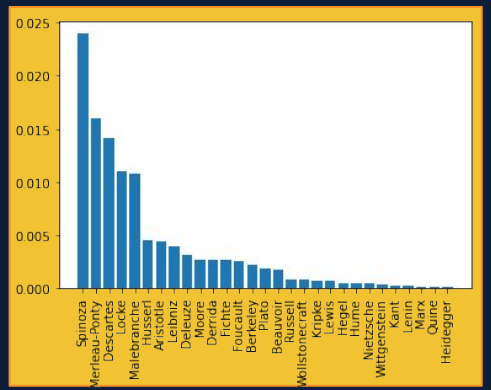


Marxism
Capitalism

Topic Assessment: Topic 2 and Topic 7

NATURE, LAWS, CAUSE,
BODY, SOUL, MIND





Rationalism

Final Topic Interpretation

German Idealism	Communism Capitalism	Rationalism
Feminism	Kantism	Empiricism Idealism
Ethics	Continental	Analytic

What to do with these topics?

- Further Evaluation
- Recommendation Systems
- Clustering
- Historical Analysis

THANKS!

Questions?