

**Análise de Sentimento  
das Principais Obras da  
Filosofia Ocidental ao  
Longo da História**

**Breno Abreu  
Pensadores de Dados**

# OBJETIVO

Compreender, a partir do uso de modelos de aprendizado de máquina para análise de linguagem, como as principais escolas de pensamento da filosofia ocidental estão correlacionadas, procurando responder às seguintes perguntas:

1. Qual é o sentimento predominante em cada escola de pensamento?
2. Quais são as conexões mais fortes entre cada escola de pensamento a partir de seus temas mais comuns?
3. Se ordenarmos as escolas em N grupos, como os grupos estariam separados?

# DADOS

O corpus utilizado para a pesquisa é o History of Philosophy, criado por Kourosh Alizadeh, ex-professor de filosofia e atual cientista de dados.

Cada linha é uma sentença de um livro escrito por um filósofo importante.

Autor | Escola | Data de Publicação | Sentença

360 000 linhas | 59 obras | 36 filósofos | 13 escolas

# MODELOS

1. **Sentiment Intensity Analyzer:** analisar cada sentença e conferir um valor entre -1 (sentimento muito negativo) e 1 (sentimento muito positivo); onde 0 indica um sentimento neutro.
2. **Empath:** apresentar a quantidade de palavras em uma sentença que estão relacionadas com cada uma das 200 categorias lexicais apresentadas pelo modelo. e.g, *hate, love, politics, confusion, animal*.
3. **K-means:** agrupar as instâncias vetorizados em N *clusters* distintos.

# EMPATH

## Empath: Understanding Topic Signals in Large-Scale Text

Ethan Fast, Binbin Chen, Michael S. Bernstein  
Stanford University  
{ethan.fast, msb}@cs.stanford.edu, bchen45@stanford.edu



Figure 2. Empath learns word embeddings from 1.8 billion words of fiction, makes a vector space from these embeddings that measures the similarity between words, uses seed terms to define and discover new words for each of its categories, and finally filters its categories using crowds.

# EMPATH

- Similar ao LWIC
- LIWC apresenta 40 categorias
- Empath apresenta 200 categorias
- O conteúdo do modelo foi validado por seres humanos
- Alta correlação entre LWIC e Empath
- Leva em consideração o contexto e associação entre termos

social media	war	violence	technology	fear	pain	hipster	contempt
facebook	attack	hurt	ipad	horror	hurt	vintage	disdain
instagram	battlefield	break	internet	paralyze	pounding	trendy	mockery
notification	soldier	bleed	download	dread	sobbing	fashion	grudging
selfie	troop	broken	wireless	scared	gasp	designer	haughty
account	army	scar	computer	tremor	torment	artsy	caustic
timeline	enemy	hurting	email	despair	groan	1950s	censure
follower	civilian	injury	virus	panic	stung	edgy	sneer

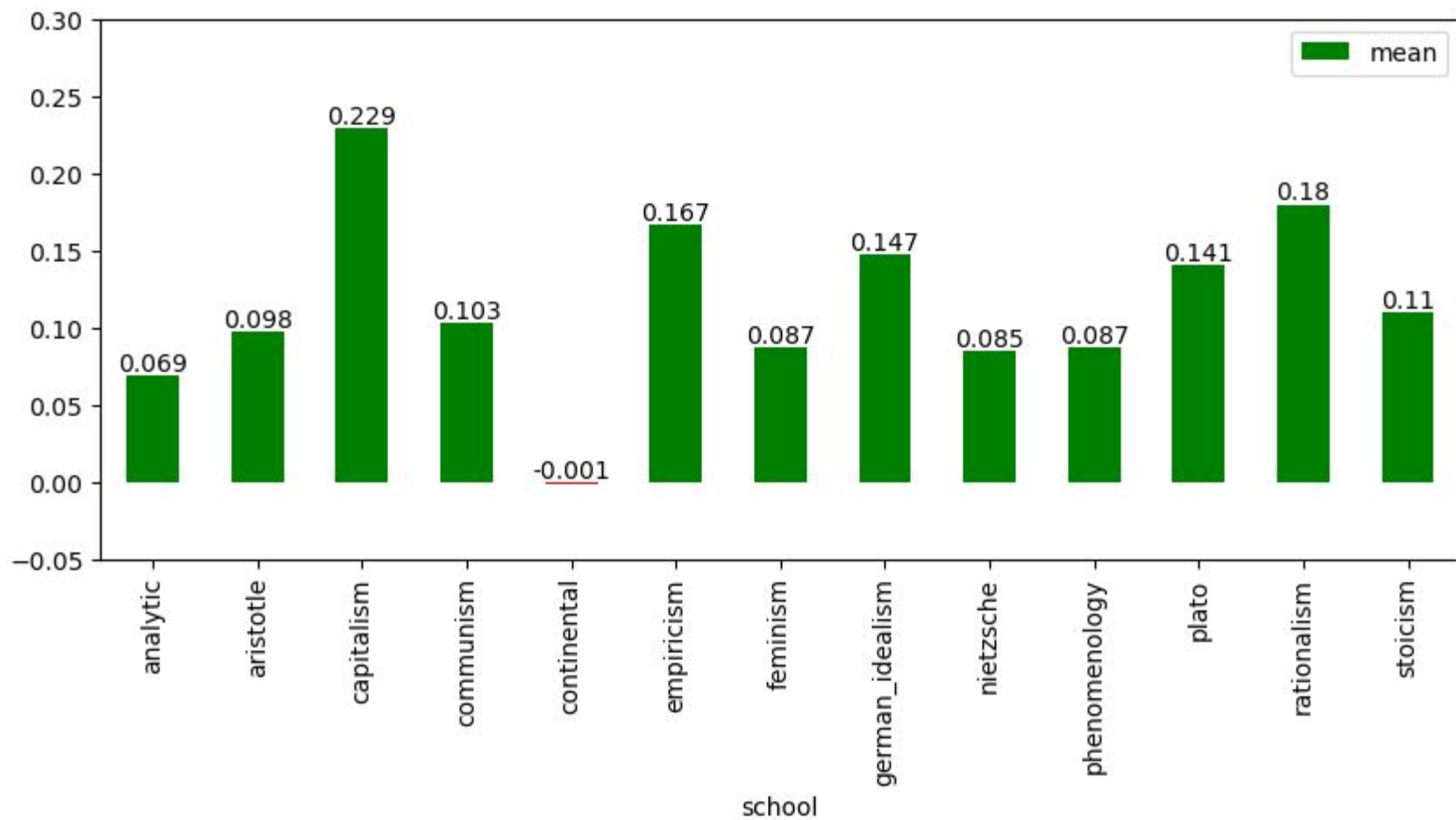
Table 1. Empath can analyze text across hundreds of data-driven categories. Here we provide a sample of representative terms in 8 sample categories.

# PERGUNTA 1

Qual é o sentimento predominante em cada escola de pensamento?

**Hipóteses:**

- A. A escola continental possui um sentimento médio mais positivo que a da escola analítica.
- B. Nietzsche apresenta uma média de sentimento negativo.



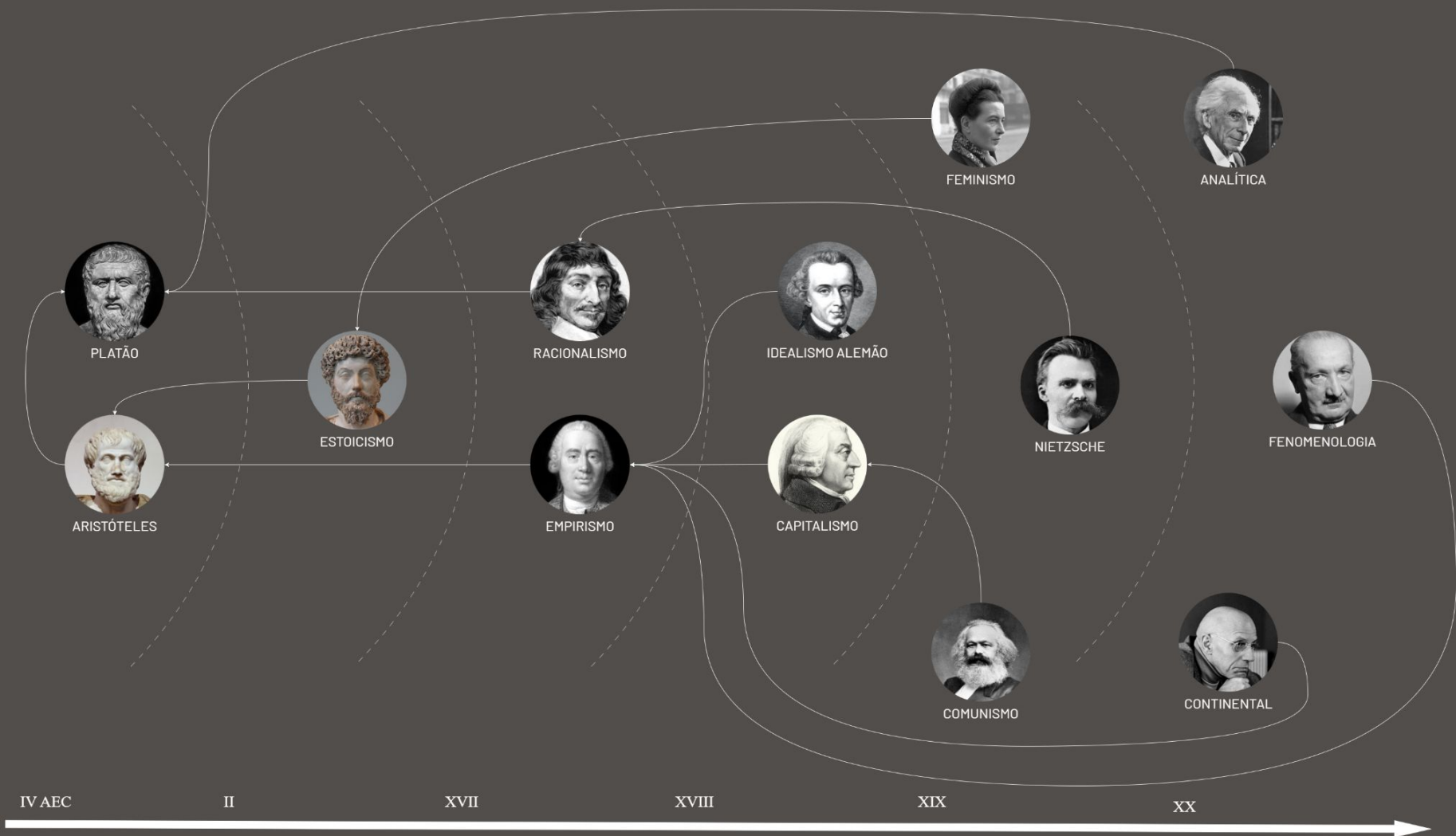


## Pergunta 2

Quais são as conexões mais fortes entre cada escola de pensamento a partir de seus temas mais comuns?

**Hipótese:**

- A. As conexões mais fortes serão com as escolas de pensamento imediatamente anteriores cronologicamente à escola em análise.





PLATÃO



ARISTÓTELES



ESTOICISMO



RACIONALISMO



EMPIRISMO



IDEALISMO ALEMÃO



CAPITALISMO



FEMINISMO



NIETZSCHE



COMUNISMO



ANALÍTICA



FENOMENOLOGIA



CONTINENTAL

IV AEC

II

XVII

XVIII

XIX

XX

## Pergunta 3

Se separarmos as escolas em N grupos, como os grupos estariam separados?

**Hipótese:**

- A. Se agruparmos as escolas de acordo com o número de séculos presente na base de dados, os grupos formados serão separados por século.



COMUNISMO



ESTOICISMO



NIETZSCHE



CAPITALISMO



FEMINISMO



RACIONALISMO



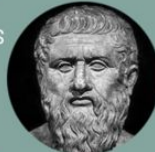
IDEALISMO ALEMÃO



EMPIRISMO



ARISTÓTELES



PLATÃO



CONTINENTAL

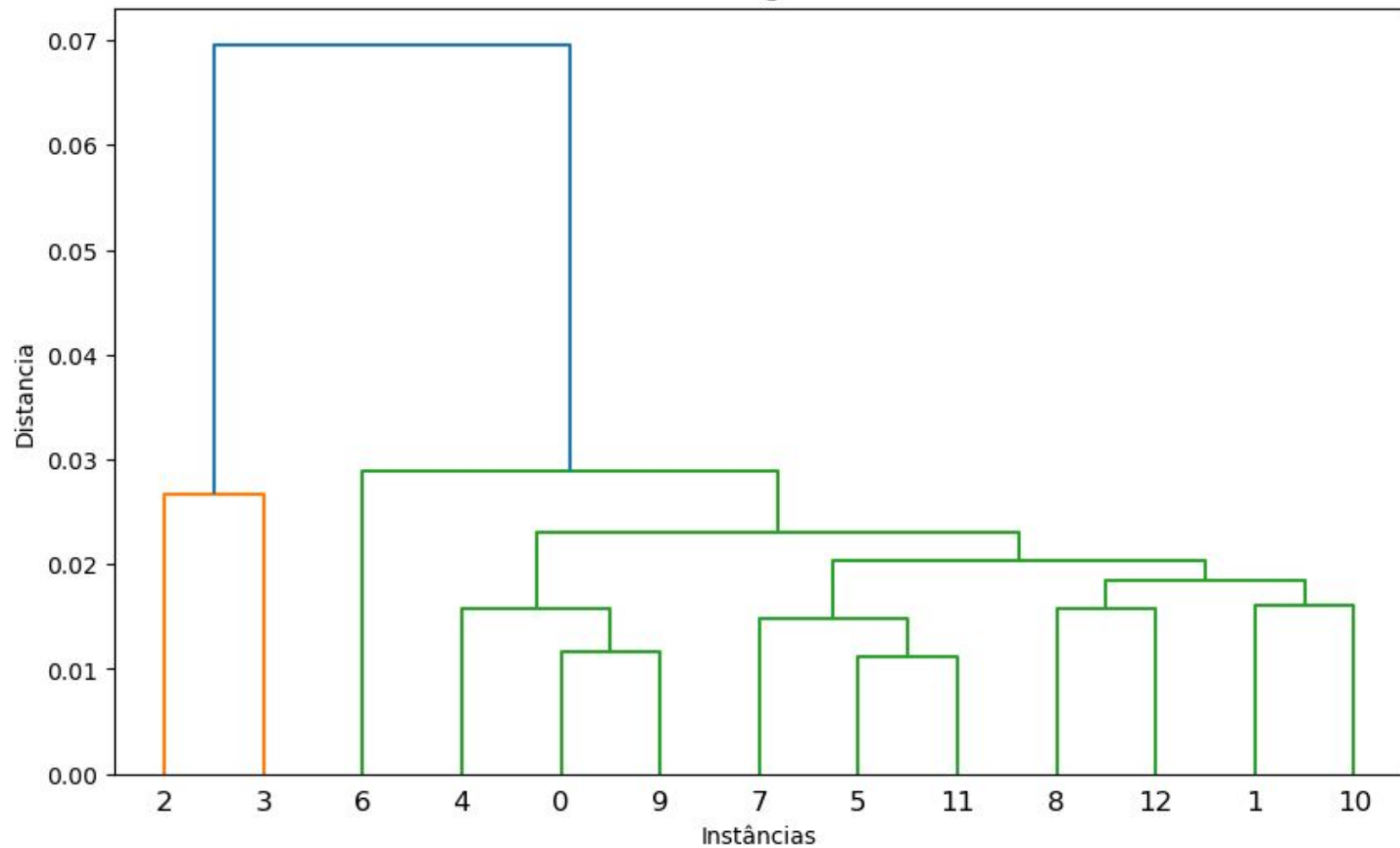


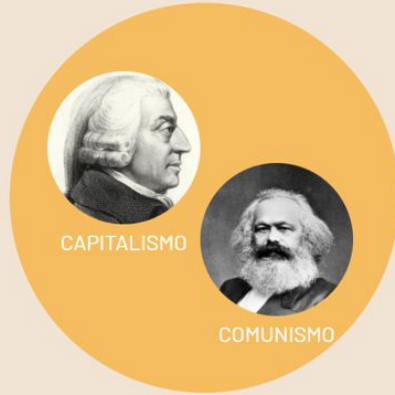
FENOMENOLOGIA



ANALÍTICA

Dendrograma





# LIMITAÇÕES E TRABALHOS FUTUROS

A principal limitação para a execução da pesquisa foi a falta de obras e autores de outras escolas de pensamento. A base, por exemplo, não inclui nenhuma obra dos períodos medieval e renascentista.

Para trabalhos futuros a proposta é criar uma nova base de dados incluindo uma variedade maior de textos assim como a inclusão de escolas de pensamento que não pertencem à tradição europeia.