Explaining Latent Dirichlet Allocation to Novice Learners

A challenge to communicate complex computation

Authors: Hannah Eli, Jaimie Murdock, Colin Allen
Indiana University, Departments of Cognitive Science and Informatics
Bloomington, IN, United States

Videos

- 3 videos planned, to explain:
  - the model through narrative
  - the model with Thomas Jefferson's memoir
  - LDA as a concept

- 1 video added later to submit to a pitch competition

3/4 of the concepts for the “LDA as a concept” video, which will be made in Adobe AfterEffects.

LDA

Latent Dirichlet Allocation (LDA) is a mixed-membership model of statistical analysis.

It also:

- is based on the work of mathematician Gustav Dirichlet. Dirichlet also worked with prime numbers, modern concepts of the function, and other important ideas in modern math.

- makes two simultaneous assumptions: documents are a distribution over topics, and topics are a distribution over words.

- can be used for document comparison and retrieval based on topic composition similarity.