Big Text Data Made Simple

Laurent El Ghaoui
Professor, EECS

CURRENT RESEARCH

Understanding and visualizing large text databases with fast machine learning

In an age of information overload, we depend on search engines like Google to scour the billions of text documents and websites to find that one video, document, or picture to suit our needs. Despite the power of search engines like Google and Yahoo, our search queries often produce convoluted search results that are difficult to browse and are often irrelevant. Dr. Laurent El Ghaoui of University of California, Berkeley focuses on developing new and efficient tools that will lessen the information overload by providing sophisticated search results based on sparse optimization and machine learning.

Dr. El Ghaoui's research is tackling information excess and data discovery. The platform StatNews.org provides a service to researchers from social sciences and humanities, allowing them to obtain a summarized image of a given topic from a large database of news, like that of the *BBC*, *New York Times* or *Wall Street Journal*. Users can view information in an interactive graph that shows key terms pertaining to the topic and their change in prevalence over time. StatNews is able to detect trends in language and media coverage of news, and develop historical perspectives. It is also able to uncover how countries, diseases or regions are portrayed in international news. The technology is extremely scalable, as it can summarize tens of thousands of documents in a few seconds on a small laptop computer, and can be applied to any text, in any language. It uses a sophisticated algorithm that gives the user the context in which a query word like "Obama" is mentioned in the news and gives users a summarized version of how the topic has evolved over time....

Read More at benefunder.com/

AFFILIATION

University of California, Berkeley

EDUCATION

- Ph.D. in Aeronautics and Astronautics 1990 , Stanford University
- M.S. in Applied Mathematics and Physics 1985 , Ecole Polytechnique, France

AWARDS

- Co-recipient of a SIAM Optimization Prize (2008)

RESEARCH AREAS

Technology, Informational Sciences / Internet

FUNDING REQUEST

Contributions to will be used to fund the StatNews.org project, which has the potential to have a powerful impact on the way news information is consumed and interpreted. Funds will support its infrastructure and the maintenance of the servers that allow for the services to be provided.