Online Semantic Analysis by Text Embedding (#7869)

A technology to analyze the semantics of unstructured text

Georgia Tech inventors have developed a method that combines delicate natural language processing methods and an unsupervised learning algorithm to extract critical and latent features (embeddings) from raw text. These features are highly separate from one another and typically have lower dimensionality and more compact representations compared to conventional language processors. In addition, the learning process does not depend on any structural or label information and can be updated by the algorithm itself with the increase of new data. The technique was originally tailored to analyze police reports, consisting of time, location, and text descriptions, but could be utilized for a variety of applications. The inventors have also developed a software interface for the text analysis algorithm.

Benefits/Advantages

- Better performance compared to the conventional natural language processing methods

Potential Commercial Applications

- Text analysis including but not limited to analysis of:
  - Insurance records
  - Medical reports
  - Crime reports
  - Social media content

Background/Context for This Invention

Sifting through unstructured text for meaning can be a labor intensive job. Currently there are algorithms to identify classifications of text and patterns within a text, but these representations can be ambiguous.

Yao Xie
Assistant Professor – Georgia Tech School of Industrial & Systems Engineering

Shixiang Zhu
Ph.D. Student – Georgia Tech School of Industrial & Systems Engineering

More Information

U.S. Application Filed - 16/383,563

Publications

Crime Incidents Embedding Using Restricted Boltzmann Machines, IEEE, September 13, 2018

For more information about this technology, please visit:
https://licensing.research.gatech.edu/technology/online-semantic-analysis-text-embedding

Images:

The automated sequential delivery of multiple fluids. A varying number of delay gates imprinted in the branches are shown in the figure.

COVID-19 and flu saliva test on paper: (A) The automatic sequential delivery of multiple reagents required for virus test; (B) Water pouring into the device triggers the virus assay, allowing the presence of SARS-CoV-2 and influenza A & B viruses to be visually identified by the color changes in the corresponding detection spot