Title:

USEVis: Visual Analytics of Attention-based Neural Embedding in Information Retrieval

Authors:

Xiaonan Ji, Washington University in St. Louis, United States; The Ohio State University, United States

Yamei Tu, The Ohio State University, United States

Wenbin He, The Ohio State University, United States

Junpeng Wang, The Ohio State University, United States

Han-Wei Shen, The Ohio State University, United States

Po-Yin Yen, Washington University in St. Louis, United States

Contact :

Xiaonan Ji ([ji.62@osu.edu](mailto:ji.62@osu.edu))

Abstract:

Neural attention-based encoders, which effectively attend sentence tokens to their associated context without being restricted by long-term distance or dependency, have demonstrated outstanding performance in embedding sentences into meaningful representations (embeddings). The Universal Sentence Encoder (USE) is one of the most well-recognized deep neural network (DNN) based solutions, which is facilitated with an attention-driven transformer architecture and has been pre-trained on a large number of sentences from the Internet. Besides the fact that USE has been widely used in many downstream applications, including information retrieval (IR), interpreting its complicated internal working mechanism remains challenging. In this work, we present a visual analytics solution towards addressing this challenge. Specifically, focused on semantics and syntactics (concepts and relations) that are critical to domain clinical IR, we designed and developed a visual analytics system, i.e., USEVis. The system investigates the power of USE in effectively extracting sentences' semantics and syntactics through exploring and interpreting how linguistic properties are captured by attentions. Furthermore, by thoroughly examining and comparing the inherent patterns of these attentions, we are able to exploit attentions to retrieve sentences/documents that have similar semantics or are closely related to a given clinical problem in IR. By collaborating with domain experts, we demonstrate use cases with inspiring findings to validate the contribution of our work and the effectiveness of our system.

Presenting Author:

Xiaonan Ji