Harnessing Scientific Literature for Boosting Discovery and Innovation

Abstract:

With millions of scientific papers coming out every year, researchers are forced to allocate their attention to increasingly narrow areas, creating isolated research bubbles that limit knowledge discovery. This fragmentation of science slows down progress and prevents the formation of cross-domain links that catalyze breakthroughs. Toward addressing these large-scale challenges for the future of science, my work develops new paradigms for searching, recommending and discovering scholarly knowledge.

In this talk, I will present methods and systems for recommending research inspirations and creating links across areas and ideas. My approach is based on extracting novel structural representations of scholarly literature, and leveraging them for “matchmaking” between problems and solutions and between scientists. I will also present a new task we introduced for jointly addressing diversity, ambiguity and hierarchy of language in scientific texts. In this new setting, the goal is to learn to construct cross-document coreference clusters along with a hierarchy defined over the clusters. As I will demonstrate, being able to automatically induce this graph structure can help unlock important applications in scientific discovery, but this remains a highly difficult open problem.