
JOHN SASTRILLO, University of British Columbia
Analyzing the Nature of Citations and Their Automatic Classification

With the ever growing number of scientific publications, navigating a field's literature for unexplored possibilities becomes correspondingly more difficult. While citations are useful in connecting the progression of ideas over the years, a common limitation persists in the lack of context for why a publication was cited. Often times, the context needs to be inferred from the text. While various platforms have attempted to present the context of citations in different ways, they often fail to give a clear definition of the nature of the citation. Instead, these platforms present the sentence where the citation is found but the sentence again lacks context without reading the entire text. In this presentation, I will discuss the pros and cons of both Machine Learning and non-Machine Learning algorithms in finding the best sentence to describe the nature of a citation and classifying the citation based on the sentence found. Furthermore, I will discuss how different platforms have used citations to facilitate researchers' understanding of their fields. With clearer definitions for why publications are cited in conjunction with an automatic way to describe the nature of a citation, we can create effective metrics for understanding the impact of a publication, and develop novel platforms for visualizing, navigating and understanding the changes in a field. On a broader level, a deeper understanding of the progression of a field overtime can highlight new areas and challenges for future research.