

Title

Entity Linking Prediction

Abstract

In the project affiliated with Worcester Polytechnic Institute (WPI) and BT called Graduate Qualifying Project (GQP), the focus will be on the entity linking aspect of Rosette. The Named-Entity Linking task aims to extract all named entities (ie. person, organization, location, etc) from a text document, called Entity Extraction/ Entity Detection/ Named-Entity Recognition, and link the identified mentions to a corresponding entity record in the Wikipedia Knowledge Base named Entity Disambiguation.

Although BT has Rosette to efficiently perform entity detection and entity linking to the database, the goal is to improve and optimize the current procedure. This would be achievable by performing entity linking along with the utilization of entity disambiguation (ED) to achieve the best possible F1 or accuracy score. In this project, we focused on the entity disambiguation task, researching and experimenting with three neural entity linking models which are End-to-End model, latent relations model and robust disambiguation model and then came up with a novel disambiguation approach. At last, we evaluated which one has the best performance to recommend to BT.