Machine learning using context vectors for object coreference resolution

Thuy Le Thi[1,2,3] · **Tuoi Phan Thi[1,2]** · **Tho Quan Thanh[1,2]**

Abstract

Object coreference resolution is used in sentiment analysis to identify sentiment words referring to an aspect of an object in a document. However, this poses a challenge in natural language processing and is consequently an area of ongoing research. Further, to the best of our knowledge, object coreference resolution with more than one object has not been given much attention. To effectively address object coreference resolution, this paper proposes a method in which machine learning is applied to a large volume of textual data represented by context vectors, constituting a new form of language representation. The proposed machine learning model uses these vectors to achieve state-of-the-art performance in object coreference resolution. In addition, a combination of dependency grammar, sentiment ontology, and coreference graphs is used to obtain triplets of object, aspect, and sentiment. In experiments conducted on sentiment textual data obtained from Amazon.com, the proposed method achieved an average coreference resolution of object, aspect, and sentiment precision value of approximately 90%. This result suggests that the proposed method can contribute considerably to the field of object coreference resolution, and further research is therefore warranted.

✉ Thuy Le Thi
  1585007@hcmut.edu.vn ; lethithuyit@iuh.edu.vn

  Tuoi Phan Thi
  tuoi@hcmut.edu.vn

  Tho Quan Thanh
  qttho@hcmut.edu.vn

1 Faculty of Computer Science & Engineering, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam

2 Vietnam National University, Ho Chi Minh City, Vietnam

3 Faculty of Information Technology, Industrial University of Ho Chi Minh City, Ho Chi Minh City, Vietnam