Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 116 (2017) 449–459

Procedia Computer Science

www.elsevier.com/locate/procedia

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017, 13-14 October 2017, Bali, Indonesia

It Takes Two To Tango: Modification of Siamese Long Short Term Memory Network with Attention Mechanism in Recognizing Argumentative Relations in Persuasive Essay

Aryo Pradipta Gema[a], Suhendro Winton[a], Theodorus David[a], Derwin Suhartono*[a], Muhsin Shodiq[a], Wikaria Gazali[b]

[a] *Computer Science Department, School of Computer Science, Bina Nusantara University, Jl. K. H. Syahdan No. 9 Kemanggisan, Jakarta 11480, Indonesia*

[b] *Mathematics Department, School of Computer Science, Bina Nusantara University, Jl. K. H. Syahdan No. 9 Kemanggisan, Jakarta 11480, Indonesia*

Abstract

We propose a novel approach in a dataset of argumentation relations. This task is intended to analyze the presence of a support relation between two sentences. To be able to identify relations between two sentences or arguments, one is obliged to understand the nuance brought by both sentences. Our models are modification of siamese network architectures, in which we replace the feature extractor into Long Short Term Memory and implement cosine distance as the energy function. Our models take a pair of sentences as their input and try to identify whether there is a support relation between those two sentences or not.The primary motivation of this research is to prove that a high degree of similarity between two sentences correlates to sentences supporting each other. This work will focus more on the modification of siamese network and the implementation of attention mechanism. Due to the difference in dataset setting, we cannot arbitrarily compare our results with the prior research results. Therefore, this work will not highlight the comparison between deep learning and traditional machine learning algorithm per se, but it will be more of an exploratory research. Our models are able to outperform the baseline score of accuracy with a margin of 17.33% (67.33%). By surpassing the baseline performance,we believe that our work can be a stepping stone for deep learning implementation in argumentation mining field.

Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

Keywords: Siamese Network; Long Short Term Memory, Attention Mechanism, GloVe, Cosine Distance

* Corresponding author. Tel.: +62-215345830
E-mail address: dsuhartono@binus.edu

1877-0509
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.
10.1016/j.procs.2017.10.036

1. Introduction

Comprehensively understand an argument or even a part of it is not an easy task. It requires the reader to understand each word that build the sentence or part of an argument itself. Not only to understand each word meaning, but also understand the relation between one word and another. Argumentation understanding does not stop only on the word level, it is more of a hierarchical problem. One should also understand the relation between two sentences. This is the fundamental idea why argumentation can actually represent the unreasonable intellectual capacity of human. An attempt to copy that ability is the foundation of argumentation mining researches.

As expected, many showed interest on argumentation mining, even though it is still a relatively young research field. Many researches forged novel approaches and achieved the state-of-the-art results. Several researches in identifying argument components relied heavily on handcrafted features[1,2]. They managed to observed several machine learning algorithms in conjunction with handcrafted features, such as contextual features and discourse marker. The notable work to identify argumentative discourse structures also relied on handcrafted features[3]. In that work, they also managed to introduce the tree of argument structure in which revolutionizes our understanding of argument structure. Taking the benefit of the tree, they introduced parse features by calculating the depth of the tree and the number of subclasses.

Deep learning algorithms have revolutionized machine learning paradigm in the last decade. Instead of spending hours to engineer features that are most suitable for a particular task, one can conveniently inject raw data to a deep learning architecture and get a comparable, if not better, results. Deep learning architectures are made ranging from Computer Vision to Natural Language Processing (NLP) domain to extract their high level features that might not be seen by human experts. A research in slot filling task proved that deep learning can perceive 42% more unseen features in comparison to handcrafted features[4].

Deep learning breakthroughs did help researches in Natural Language Processing (NLP) tasks. The invention of word vector representation[5,6] reformed the way researchers quantify words. The development of Recurrent Neural Network (RNN)[7] and its modifications[8,9] and Convolutional Neural Network (CNN) in sentence classification[10] are the state-of-the-art algorithms in Natural Language Processing (NLP). Several recent argumentation mining researches are attempted to follow the deep learning trend as well. In 2016, an attempt[11] used Bidirectional Long Short Term Memory architecture to compare the level of convincingness of two arguments. A more recent work[12], implemented CNN to recognize insufficiently supported arguments. Another work focused on the implementation of word vector representation and Long Short Term Memory unit to identify argument components[13].

This paper attempted to focus more on the exploration of deep learning in a task which dataset is released and havent tried yet. We adopt the latest publicly available dataset of argumentation structure[14] which consists of 402 essays and do several experiments on the argument relations task. As the highlight, we implemented Long Short Term Memory (LSTM) siamese network with cosine proximity as the energy function to identify argumentative relations between pairs of sentences. We argue that when a sentence supports another sentence, there must be a certain degree of similarity between both sentences and we believe that cosine proximity can represent that similarity metric better than other reasonable alternatives. In this paper, we presented results of our model that can outperform the baseline of this research. As a disclaimer, this will not be comparison experiments towards the prior research results[1], considering that we used a different dataset. Instead, this is more of an analysis of the deep learning algorithm in argumentation mining field.

2. Related Works

Identifying argumentative relations is a part of argumentation mining tasks. In order for one to identify an argumentative relation between two sentences, one needs to comprehensively understand the semantic substance of both sentences. That's why this task can be classified as an argument analysis subdomain. Extensive researches in argument analysis subdomain are exponentially increasing nowadays. For example, automatic argumentation summarizer in a political debate corpus[15] which requires an understanding of the entire debate implicit meaning. An attempt to automatic convincingness level prediction and ranking by using a pairwise learning process Support Vector Machine (SVM) and Bidirectional Long Short Term Memory (BiLSTM)[11]. This task requires the machine to understand the nuance of each argument before it can decide which argument is more convincing.