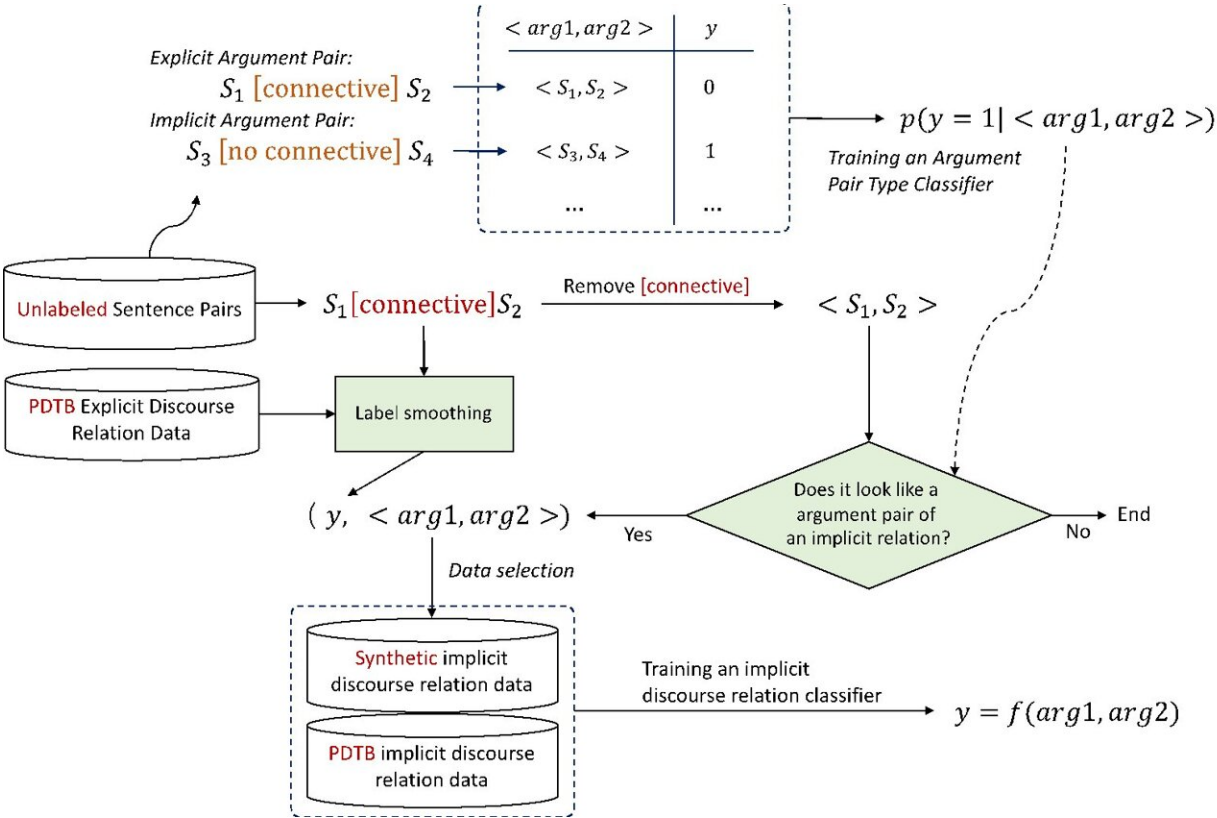


A new and simple approach extending data expansion-based implicit discourse relation classification

August 27 2024



The processing flow of the proposed method. Credit: *Frontiers of Computer Science* (2023). DOI: 10.1007/s11704-023-3058-2

Discourse relation classification is a fundamental task for discourse

analysis, which is essential for understanding the structure and connection of texts. Implicit discourse relation classification aims to determine the relationship between adjacent sentences and is the most challenging in discourse relation classification because it lacks explicit discourse connectives as linguistic cues and sufficient annotated training data.

A promising way is to expand the [training data](#) for implicit discourse relations based on easy-to-collect explicit discourse relations. However, the expanded data often involves noise in both the argument pair selection and discourse relation [sense](#) assignment process, leading to limited improvements.

A team led by Wei Song propose a novel method for explicit data expansion to address the above challenges. They published their [research](#) on 15 August 2024 in *Frontiers of Computer Science*.

In order to obtain suitable argument pairs, the argument pair type classification (APTC) task is proposed. APTC is a classifier that can distinguish between explicit and implicit argument pairs, and select only those explicit argument pairs that are similar to natural implicit argument pairs for data expansion, so the method can filter out noisy and unsuitable argument pairs for data expansion.

To annotate the senses of expanded argument pairs, a simple label-smoothing strategy is proposed. Instead of assigning a single dominant sense to a discourse connective, a smoothed sense is derived based on the distribution of each sense for that discourse connective.

This can reduce the impact of noisy sense labels that may not match the actual relation between the arguments.

Despite its simplicity, the evaluation results on PDTB 2.0 and PDTB 3.0

demonstrate the effectiveness of the proposed method. It can consistently lead to improvements compared with previous data expansion methods and obtain competitive performance to the state-of-the-art models across datasets and on both the top-level class and the second-level type senses.

The discriminative explicit argument pair selection and the label-smoothing strategy complement and depend on each other to achieve the best performance. The results and analysis confirm that the proposed method extends the data expansion-based implicit discourse relation [classification](#).

More information: Wei Song et al, Discriminative explicit instance selection for implicit discourse relation classification, *Frontiers of Computer Science* (2023). [DOI: 10.1007/s11704-023-3058-2](https://doi.org/10.1007/s11704-023-3058-2)

Provided by Frontiers Journals

Citation: A new and simple approach extending data expansion-based implicit discourse relation classification (2024, August 27) retrieved 3 September 2024 from <https://techxplore.com/news/2024-08-simple-approach-expansion-based-implicit.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.