

INTISARI

IDENTIFIKASI KALIMAT KUTIPAN DARI TEKS BERITA *ONLINE* BERBAHASA INDONESIA DENGAN METODE BERBASIS ATURAN

Oleh

YUSUF SYAIFUDIN

11/313303/PA/13672

Berita merupakan suatu sumber informasi yang sering digunakan sebagai data dalam ekstraksi informasi. Pernyataan atau opini narasumber merupakan salah satu informasi dalam berita yang cukup penting untuk mengetahui hal-hal yang dikatakan oleh seseorang atau organisasi terkait peristiwa tertentu. Pernyataan narasumber ini menjadi penting karena merupakan salah satu informasi yang dapat mempengaruhi pikiran pembaca berita. Selain itu, pernyataan tersebut juga dapat menunjukkan sentimen pembicara terkait suatu kejadian yang diberitakan. Pernyataan tersebut sering ditulis menggunakan metode kutipan, baik kaidah penulisan kutipan langsung maupun tidak langsung.

Penelitian ini menggunakan pendekatan metode berbasis aturan (*rule-based method*) di mana kutipan-kutipan dalam berita diekstrak berdasarkan kaidah penulisan kutipan yang umum terdapat pada berita. Setiap pola penulisan kutipan yang ditemukan dalam kumpulan data berita dibuat menjadi satu kelas aturan tersendiri yang mewakili pola tersebut. Fitur yang digunakan pada pembuatan kelas aturan ialah keberadaan *reporting verb* dan entitas dalam suatu kalimat.

Hasil penelitian ini menunjukkan bahwa seluruh kelas aturan yang dibuat telah mampu mengekstraksi kalimat pernyataan dengan cukup baik. Sistem mampu mendeteksi kalimat pernyataan dengan akurasi sebesar 88,618%, presisi sebesar 99,013% dan *recall* sebesar 79,936%.

Kata kunci: ekstraksi kutipan, ekstraksi opini, *rule-based*, ekstraksi informasi.

ABSTRACT

QUOTATIONS IDENTIFICATION FROM INDONESIAN ONLINE NEWS USING RULE-BASED METHOD

By

YUSUF SYAIFUDIN

11/313303/PA/13672

News is a source of information that is often used as data in information extraction. Interviewees statement or opinion is one of the sources of information in the news, which is quite important to know things that are said by a person or organization related to a specific event. Interviewees statement is important because it is one of the information that can affect the mind of a newsreader. Moreover, the statement may also indicate the speaker's sentiment related to an event that was reported. The statement is often written using quotation, by writing rules of either direct quotes or indirect quotes.

This study uses a rule-based method which quotes in the written news are extracted based on common rules contained in the news. Each pattern of writing quotations found in a collection of news data is made into a class of rule that represents the pattern. Features used in the class of rules is the existence of reporting verb and entities in a sentence.

The result of this study indicate that the entire class rules have been made are able to extract sentence statement pretty well. The system is able to detect sentence statement with an accuracy of 88.618%, precision of 99.013% and recall 79.936%.

Key words: quotations extraction, opinion extraction, rule-based, information extraction.