

Abstraksi

Pengklasifikasian e-mail apakah e-mail tersebut adalah suatu spam atau ham akan sangat membantu kita dalam menyaring e-mail yang masuk ke inbox kita untuk menyingkirkan e-mail yang tidak diinginkan (spam). Hal tersebut tentu saja perlu dilakukan secara otomatis berdasarkan aturan atau rule yang kita inginkan. Namun dengan semakin berkembangnya teknik spamming memaksa kita untuk meningkatkan variasi rule tersebut dan tentu saja penentuan rule pun perlu dilakukan secara otomatis.

Oleh karena itu diperlukan suatu metode pembelajaran tentang mekanisme pembangkitan rule secara otomatis berdasarkan data-data yang sudah ada. Teknik ini dapat menemukan rule secara otomatis dan diharapkan dapat berlaku umum untuk data-data yang belum diketahui. Salah satu teknik yang dapat dipakai adalah dengan metode yang sesuai dengan konsep Cyc Ontology.

Dengan konsep Cyc Ontology dapat dihasilkan suatu fungsi atau program yang dapat menghasilkan suatu rule berdasarkan kemiripan nilai semantik dari suatu data dalam e-mail dengan data-data yang sudah ada. Hal ini yang menjadi acuan dalam proses klasifikasi.

Dalam tugas akhir ini, dibuat perangkat lunak untuk mengimplementasikan konsep Cyc Ontology untuk proses pembangkitan rule dan klasifikasi pada spam filtering. Hasil yang didapat menunjukkan bahwa penentuan rule yang lebih kompleks memberikan akurasi yang lebih tinggi.

Kata kunci: spam filtering, rule, klasifikasi, cyc, ontologi.