

ABSTRAK

Mendapatkan representasi argumen semantik dari sebuah kalimat adalah hal yang penting dalam *text mining*, seperti *information extraction* dan *question answering*. Dalam *semantic role labelling* pemilihan fitur menjadi hal yang berpengaruh pada kinerjanya, dengan kata lain juga berpengaruh pada *recall* dan *precision* yang dihasilkan. Permasalahan saat ini adalah bagaimana cara mengkombinasikan fitur dan kombinasi seperti apa yang digunakan agar pelabelan peran semantik mendapatkan kinerja yang diharapkan. Tugas akhir ini mencoba untuk menganalisis dan mengkombinasikan beberapa fitur yang sudah ada dan terbukti telah menunjukkan hasil performa yang optimal pada penelitian sebelumnya. Fitur yang digunakan pada tugas akhir ini adalah *BaselineFeature* ditambah *Noun Head of PP*, *First Word in Constituent*, *Syntactic Frame*, *Argument Order*, dan *Constituent Order* yang nantinya akan digunakan dalam klasifikasi argumen semantik. Hasil dari penelitian ini menunjukkan bahwa tidak semua kombinasi fitur secara acak dapat meningkatkan performa klasifikasi argumen semantik, ada beberapa fitur yang justru akan menurunkan performa klasifikasi bila tidak dikombinasikan dengan fitur yang tepat. Berdasarkan rata-rata seluruh skenario, kombinasi terbaik adalah kombinasi dengan penggunaan kelima fitur tambahan yang digunakan dalam penelitian ini dengan akurasi sebesar 75,34% dan *f-score* sebesar 74,6%. Penambahan *data training* juga dapat meningkatkan performa klasifikasi.

Katakunci : Kombinasi Fitur, Pelabelan peran semantik, Klasifikasi

ABSTRAK

Mendapatkan representasi argumen semantik dari sebuah kalimat adalah hal yang penting dalam *text mining*, seperti *information extraction* dan *question answering*. Dalam *semantic role labelling* pemilihan fitur menjadi hal yang berpengaruh pada kinerjanya, dengan kata lain juga berpengaruh pada *recall* dan *precision* yang dihasilkan. Permasalahan saat ini adalah bagaimana cara mengkombinasikan fitur dan kombinasi seperti apa yang digunakan agar pelabelan peran semantik mendapatkan kinerja yang diharapkan. Tugas akhir ini mencoba untuk menganalisis dan mengkombinasikan beberapa fitur yang sudah ada dan terbukti telah menunjukkan hasil performa yang optimal pada penelitian sebelumnya. Fitur yang digunakan pada tugas akhir ini adalah *BaselineFeature* ditambah *Noun Head of PP*, *First Word in Constituent*, *Syntactic Frame*, *Argument Order*, dan *Constituent Order* yang nantinya akan digunakan dalam klasifikasi argumen semantik. Hasil dari penelitian ini menunjukkan bahwa tidak semua kombinasi fitur secara acak dapat meningkatkan performa klasifikasi argumen semantik, ada beberapa fitur yang justru akan menurunkan performa klasifikasi bila tidak dikombinasikan dengan fitur yang tepat. Berdasarkan rata-rata seluruh skenario, kombinasi terbaik adalah kombinasi dengan penggunaan kelima fitur tambahan yang digunakan dalam penelitian ini dengan akurasi sebesar 75,34% dan *f-score* sebesar 74,6%. Penambahan *data training* juga dapat meningkatkan performa klasifikasi.

Katakunci : Kombinasi Fitur, Pelabelan peran semantik, Klasifikasi

ABSTRAK

Mendapatkan representasi argumen semantik dari sebuah kalimat adalah hal yang penting dalam *text mining*, seperti *information extraction* dan *question answering*. Dalam *semantic role labelling* pemilihan fitur menjadi hal yang berpengaruh pada kinerjanya, dengan kata lain juga berpengaruh pada *recall* dan *precision* yang dihasilkan. Permasalahan saat ini adalah bagaimana cara mengkombinasikan fitur dan kombinasi seperti apa yang digunakan agar pelabelan peran semantik mendapatkan kinerja yang diharapkan. Tugas akhir ini mencoba untuk menganalisis dan mengkombinasikan beberapa fitur yang sudah ada dan terbukti telah menunjukkan hasil performa yang optimal pada penelitian sebelumnya. Fitur yang digunakan pada tugas akhir ini adalah *BaselineFeature* ditambah *Noun Head of PP*, *First Word in Constituent*, *Syntactic Frame*, *Argument Order*, dan *Constituent Order* yang nantinya akan digunakan dalam klasifikasi argumen semantik. Hasil dari penelitian ini menunjukkan bahwa tidak semua kombinasi fitur secara acak dapat meningkatkan performa klasifikasi argumen semantik, ada beberapa fitur yang justru akan menurunkan performa klasifikasi bila tidak dikombinasikan dengan fitur yang tepat. Berdasarkan rata-rata seluruh skenario, kombinasi terbaik adalah kombinasi dengan penggunaan kelima fitur tambahan yang digunakan dalam penelitian ini dengan akurasi sebesar 75,34% dan *f-score* sebesar 74,6%. Penambahan *data training* juga dapat meningkatkan performa klasifikasi.

Katakunci : Kombinasi Fitur, Pelabelan peran semantik, Klasifikasi