

Penggunaan model ruang vektor dalam proses pengenalan relasi antar entitas pada sistem ekstraksi informasi

Kiki Marjuki, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=124259&lokasi=lokal>

Abstrak

Sistem Ekstraksi Informasi merupakan rangkaian proses untuk mendapatkan informasi dari suatu teks. Salah satu proses dalam Sistem Ekstraksi Informasi adalah Template Relation. Template relation bertugas mengenali relasi antar entitas bernama seperti person, organization, position, dan location. Penelitian ini bertujuan mengembangkan sistem pengenalan relasi antara dua entitas (person-organization, organization-position, position-person) dan relasi antara tiga entitas (personorganization- position). Pendekatan yang digunakan adalah machine learning dengan model ruang vektor. Pada pendekatan ini terdapat dua proses, yaitu proses training untuk pelatihan sistem dan proses testing untuk pengenalan sistem. Proses training menghasilkan pola-pola relasi untuk digunakan sistem pada tahap testing. Pada proses testing dicari nilai kesamaan (similarity coefficient) dari pasangan entitas pada dokumen testing dengan pola-pola relasi yang dihasilkan sebelumnya. Suatu pasangan entitas bernama dikatakan berhubungan atau berelasi jika nilai kesamaannya berada di atas nilai ambang (threshold). Uji coba menggunakan 120 dokumen yang berasal dari media masa online berbahasa Indonesia (www.kompas.com dan www.republika.co.id). 60 dokumen diantaranya digunakan sebagai dokumen training dan sisanya digunakan sebagai dokumen testing. Hasil uji coba sistem dapat mengenali relasi antar entitas pada teks dokumen dengan nilai F-measure sebesar 76,2% untuk relasi person-organization , 84,5% untuk relasi organization-position, 74,5% untuk relasi position-person, dan Fmeasure sebesar 74,9% pada relasi person-organization-position.