

Abstrakt:

Cílem naší práce byla tvorba slovníků subjektivity pro indonéštinu překladem již existujících anglických slovníků a jejich následnou kombinací průnikem a sjednocením. Porovnali jsme efektivnost výsledných slovníků pomocí jednoduché prediktivní metody, která měří a porovnává počet výskytů kladných a záporných výrazů ve větě. Použili jsme také dvě různé hodnotící funkce založené jak na četnosti tak na relativní četnosti výrazů v neanotovaných datech. Úpravou prediktivní metody využívající strojového učení jsme posléze lépe začlenili údaje, které nemohly být zachyceny jednoduchou predikcí. Dále jsme ukázali, že slovníky byly v predikci na testovacích větách schopny dosáhnout vysokého pokrytí ale nízké přesnosti. Ohodnocování výrazů dokáže zlepšit pokrytí či přesnost, ovšem však vždy i se srovnatelným poklesem v druhé metrice. Predikce na bázi strojového učení byla schopna minimalizovat citlivost výkonu na velikosti slovníku, bude však zapotřebí dalších experimentů, aby se našla nejlepší volba prediktivní metody.