

Abstrak

Keterkaitan semantik adalah salah satu jenis pengukuran yang ada pada *text mining* untuk menggambarkan bagaimana hubungan antara kata. Tujuan dari pengukuran keterkaitan semantik ini adalah untuk memperoleh nilai yang merepresentasikan seberapa besar keterkaitannya. *Pointwise Mutual Information* (PMI) merupakan salah satu pengukuran secara statistik untuk keterkaitan semantik yang telah banyak digunakan. Penerapan PMI diketahui mengalami bias untuk sepasang kata dengan frekuensi rendah, hal ini menyebabkan adanya pengembangan berupa varian pada pengukuran PMI untuk menghindari keadaan bias tersebut. *Positive Pointwise Mutual Information Cosine* (PPMIC) merupakan salah satu varian yang digunakan dalam tugas akhir ini untuk menghitung keterkaitan semantik. Perhitungan nilai PPMIC dilakukan pada dataset yang didapat dari *brown corpus*. Nilai PPMIC yang didapat oleh aplikasi dihitung korelasinya dengan Word-Sim-353 yang merupakan indeks keterkaitan kata berdasarkan sudut pandang manusia. Hasil dari penelitian pada tugas akhir ini merupakan nilai korelasi antara skor yang dihasilkan sistem dengan *gold standard SimLex-999, WordSim353 dan Miller and Charles* yang akan menghasilkan nilai kolerasi yang akan menunjukkan seberapa akurat metode pengukuran PPMIC.

Kata Kunci: Keterkaitan Semantik, *Pointwise Mutual Information, Positive Pointwise Mutual Information Cosine*.