# **BAB II**

**Kajian Pustaka**

## **Kajian Terdahulu**

Pada bagian ini akan dipaparkan mengenai kajian penelitian terdahulu yang akan dijadikan sebagai acuan dasar dalam penelitian. Selain kajian terdahulu, akan dijelaskan tentang kajian teori yang mendukung penelitian ini. Berikut beberapa kajian terdahulu yang menjadi acuan penelitian:

Nurul Justina Mahardianing Verdaningroem dan Aries Saifudin. (2018) memaparkan penelitian terkait tentang penerapan kamus dasar pada algoritma Porter untuk mengurangi kesalahan stemming Bahasa Indonesia. Penelitian ini bertujuan untuk membandingkan ketepatan algoritma Nazief & Adriani dengan kemampuan Algoritma Porter, proses pembandingan Algoritma porter dan Nazief & Adriani dilakukan dengan membuat program sederhana yang memproses dokumen teks Bahasa Indonesia sehingga diketahui stem, waktu proses, presisi dari hasil *stemming* dokumen tersebut. Pengujian Algoritma dilakukan pada 30 dokumen teks Bahasa Indonesia dengan ukuran dokumen yang bervariasi. Penelitian ini melakukan eksperimen menggunakan komputer untuk melakukan proses *stemming*. Sedangkan perangkat lunak yang digunakan untuk menyimpan *database* kamus adalah MySQL.

**Gambar 2.1** Hasil Pengujian 30 Kata Yang Dipilih Secara Acak



Dari Gambar 2.1 hasil pengujian di atas menyatakan bahwa ada beberapa kata imbuhan yang di *stemming* dengan kamus menghasilkan kata dasar yang rata-rata masih memiliki makna, namun pada hasil kata dasar yang di stemming tanpa kamus menghasilkan beberapa kata dasar yang ambigu atau membingungkan seperti “gum”, ”rogram”, “gala”, “njara”, “iksa”, ”masu” dan “ail”. Sementara untuk perbandingan waktu menunjukkan bahwa waktu yang diperlukan dalam proses *stemming* dengan kamus membutuhkan waktu beberapa detik lebih lama dari pada *stemming* tanpa menggunakan kamus. Dapat dilihat pada kolom waktu *stemming*, pada *stemming* tanpa kamus menunjukkan bahwa dari 30 kata berimbuhan yang diproses stemming rata-rata menunjukkan waktu 0,00030 detik, sementara untuk *stemming* dengan menggunakan kamus membutuhkan lebih dari 0,001 detik pada Gambar 2.1 tersebut menunjukkan bahwa rata-rata waktu *stemming* dengan kamus adalah 0,50767 detik. Percobaan waktu ini dicoba pada proses *stemming* setiap 1 kata berimbuhan. Hasil dari proses *stemming* dengan memasukkan 30 kata secara langsung dan memprosesnya secara bersamaan menghasilkan waktu 0,00813 detik untuk proses *stemming* tanpa kamus dan 15,50625 detik untuk proses *stemming* dengan menggunakan kamus. Selisih waktu yang dihasilkan adalah 15,49812 detik. Hasil dari persentase kata dasar yang di *stemming* dengan kamus dapat dilihat pada perhitungan berikut: ((30 - 5)) / 30 X 100% = 83.333%. Sementara hasil kata dasar yang di *stemming* tanpa kamus dapat dilihat pada perhitungan berikut: ((30 - 9)) / 30 X 100% = 70%. Dari hasil presentase menunjukkan bahwa proses *stemming* dengan ditambahkan kamus dasar memiliki persentase keakuratan kata yang benar lebih besar 13.333% dari pada persentase keakuratan *stemming* yang hanya menggunakan aturan yang sudah ada.

Selanjutnya dalam penelitian Wahyudi dkk. (2017) memaparkan tentang implementasi dan analisis Algoritma *stemming* Nazief & Adriani dan Porter pada dokumen berbahasa Indonesia. Penelitian ini bertujuan untuk