Abstrak

Data mining merupakan proses untuk memperoleh informasi pada suatu data, sehingga dapat diambil manfaat dari data tersebut. Terdapat beberapa proses dalam data mining, salah satunya klasifikasi. Voting feature intervals (VFI) merupakan salah satu classifier pada data mining. Model pada VFI dibangun dengan interval dan kelas. Interval pada VFI dibentuk di setiap atribut berdasarkan nilai atribut itu sendiri. Setelah terbentuk interval, kelas dihitung pada tiap-tiap interval.

Grafem to fonem atau Grapheme to Phoneme (G2P), adalah data kata dan cara baca dari kata tersebut. Dengan adanya model, maka dapat dilakukan prediksi cara baca suatu kata (mencari nilai fonem). Dataset diambil dari Kamus Besar Bahasa Indonesia (KBBI) ditambah dengan kata sehari-hari, sehingga software diharapkan dapat belajar kata-kata berimbuhan.

Berdasarkan hasil penelitian pada Tugas Akhir ini, VFI mempunyai waktu tercepat dalam pembentukan model dan nilai akurasi yang cukup tinggi dibanding dengan metode yang lain.

Kata kunci : *mining*, *interval*, VFI, G2P, waktu, akurasi