

ABSTRAK

Pada permasalahan dimensi tinggi dan banyaknya data, terdapat metode untuk mempermudah dalam pencarian informasi ataupun kategorisasi yaitu *Document Classification*/klasifikasi dokumen. Klasifikasi dokumen sendiri memiliki banyak metode, salah satunya yang akan dibahas pada Tugas Akhir ini adalah klasifikasi dokumen dengan metode kNN (*K-Nearest Neighbor*). kNN merupakan salah satu metode yang populer digunakan pada klasifikasi dokumen dikarenakan metodenya mudah untuk diimplementasi. kNN memiliki kekurangan yaitu pada saat klasifikasi, menggunakan nilai k untuk mengambil sebanyak k terdekat, dan dibutuhkan waktu untuk mencari nilai k untuk hasil optimal. Terdapat kasus pada dataset dengan persebaran data yang tidak seimbang, yaitu kelas dokumen yang terklasifikasi dengan nilai k yang kecil, yaitu kelas yang memiliki distribusi kelas lebih rendah pada *training set*, jika menggunakan nilai k yang besar, maka tetangga yang diambil akan lebih banyak tetangga yang merupakan kelas yang distribusi kelasnya lebih besar, dan mengakibatkan klasifikasi bisa menjadi kurang akurat. Metode k dinamis ini dipakai untuk menentukan pengambilan k -tetangga yang tepat pada klasifikasi kNN berdasarkan distribusi kelas sehingga dapat membuat klasifikasi menjadi lebih baik.

Kata Kunci : Document Classification, kNN, Dynamic kNN, Improved kNN, *imbalance class*