

## Abstrak

Banyaknya jumlah informasi yang berkembang baik secara online maupun offline menyebabkan pengguna mengalami kesulitan menemukan informasi yang penting. Pada umumnya informasi tersebut tidak terstruktur. Hal ini mendorong kebutuhan untuk memperoleh informasi yang bermanfaat. Salah satu cara pengelolaan informasi ini yaitu kategorisasi teks yang dapat membantu pengguna memperoleh informasi yang akurat dan tepat.

Dalam kategorisasi teks biasanya menggunakan referensi yang sama dalam pengelompokan katanya yaitu seberapa sering kata tersebut muncul di dalam dokumen, untuk itu digunakanlah pendekatan baru yaitu *Distributional Features* yang mengambil beberapa pertimbangan dalam pengelompokan dokumen seperti *compactness of the appearances* dan *position of the first appearance* dari sebuah kata. *Distributional features* dapat membantu meningkatkan performansi dengan tambahan biaya komputasi yang sedikit.

Dalam tugas akhir ini menggunakan beberapa pembobotan seperti TFIDF dan LOGTF yang dikombinasikan dengan *Distributional Features*. Untuk membuktikan efek dari *Distributional features* ini digunakan metode klasifikasi *k-nearest neighbor* (kNN).

**Kata kunci:** *distributional features*, TFIDF, LOGTF, kNN