

## Abstrak

*Folksonomy* merupakan metode kategorisasi dokumen yang tidak hierarkis, menyamaratakan kedudukan setiap kategori, dan judul kategori ditentukan secara bebas oleh siapa saja yang memasukkan sebuah dokumen di dalam kategori-kategori tersebut. Pembuatan kategorisasi dilakukan secara otomatis pada saat dokumen dimasukkan, yaitu dengan cara mengetikkan daftar kategori yang kira-kira cocok untuk dokumen tersebut. Situs [del.icio.us](http://del.icio.us) (<http://del.icio.us>) merupakan salah satu *social bookmarking site* terpopuler yang menggunakan *folksonomy*.

Penggunaan *folksonomy*, meski sangat mudah, juga mempunyai beberapa kelemahan, yaitu penggunaan *tag* yang berbeda-beda untuk konsep yang sama, penggunaan *tag* yang sama untuk konsep yang berbeda-beda, tidak adanya pengendalian mutu, dan lain-lain. Di sini penulis mencoba memberikan solusi untuk sebagian masalah tersebut yaitu dengan cara menganalisa isi dari dokumen Web yang ditunjuk dan mengkategorisasikan link tersebut secara otomatis ke beberapa *tag* menggunakan multinomial naive Bayes classifier.

Bayes classifier bekerja berdasarkan sekumpulan bukti (*evidence*) dan kelas (*class*). Dengan melakukan pelatihan (*training*) terhadap sebagian data sampel, dapat ditentukan probabilitas kepastian (*likelihood probability*) dari sebuah bukti jika diberikan kelas tertentu. Bayes classifier juga menggunakan probabilitas sebelumnya (*prior probability*) dari sebuah kelas, yang perhitungannya dapat didasarkan dari sampel data tersebut. Dari analisa sampel data tersebut, jika diberikan sebuah dokumen baru yang terdiri dari sekumpulan bukti, probabilitas setiap kelas terhadap dokumen tersebut (*posterior probability*) dapat ditentukan.

Sistem ini diimplementasikan menggunakan PHP 5, Apache, dan MySQL. Kesimpulan yang didapatkan dari penelitian ini adalah metode Bayes dapat digunakan untuk melakukan kategorisasi dokumen secara otomatis maupun sebagai alat bantu untuk kategorisasi manual.

**Kata kunci:** naive Bayes, text classification, folksonomy, indexing