

ABSTRAK

Semakin padatnya informasi pada *world wide web (www)* membawa pada permasalahan yang sulit dalam pencarian informasi yang benar-benar terfokus pada kebutuhan user dalam pencarian informasi, Karena setiap waktu mengalami perkembangan dalam *pengupdatean* sebuah informasi. Pertumbuhan dan perubahan dalam pencarian yang sesuai pada sebuah informasi, baik itu informasi yang terkini maupun informasi yang sudah lama makin diminati masyarakat dunia. Menurut penelitian yang dilakukan *dmoz open directory project* jumlah situs bahasa indonesia sampai pada tahun 2011 berkisar 2.684 situs link. Oleh sebab itu web *crawler* merupakan salah satu komponen dalam mesin pencari dan digunakan untuk mencari kesemua halaman-halaman internet untuk membuat *index* dari data yang dicarinya sehingga dapat menghasilkan data-data yang lebih akurat dengan langkah yang efisien. Selain *crawler* terdapat juga *classifier* untuk membangun sebuah *focused crawler*. Dalam pengklasifikasian ini menggunakan *naïve bayes classifier* yang merupakan metode klasifikasi. Setelah melakukan uji coba dari hasil jumlah klasifikasi berbanding dari jumlah dokumen ujicoba sebanyak 10844 menghasilkan keakuratan rata-rata berkisar 12%.

Kata kunci : Informasi Pariwisata, *Crawler*, *Classifier*, *Naïve Bayes Classifier*.