

Abstrak

Twitter adalah salah satu media sosial yang sangat populer di dunia. Hampir setiap orang memanfaatkan fitur *tweet* yang ada pada media sosial ini untuk bertukar informasi dengan orang disekitarnya. Fitur tersebut bersifat *broadcast* sehingga dapat memudahkan pengguna untuk menyebarkan informasi kepada pengguna lainnya. Seiring berkembangnya kebutuhan masyarakat, sekarang ini banyak orang menggunakan fitur *tweet* tersebut untuk kebutuhan bisnis seperti pengiklanan, promosi produk, dan lain-lain. Meskipun informasi dalam satu *tweet* hanya terbatas 140 karakter, biasanya mereka para pengguna yang ingin mempromosikan sesuatu di *Twitter* melampirkan *source link* kedalam *tweet* nya untuk memberikan keterangan lebih lanjut tentang isi dari *tweet* promosi tersebut. Tentu saja informasi tersebut sangat bermanfaat bagi mereka para pengguna lainnya yang menginginkan informasi tersebut, namun sebaliknya ini akan menjadi hal yang merugikan dan bahkan dapat dianggap sebagai *spam* apabila isi dalam *tweet* promosi tersebut tidak relevan dengan link nya. Dengan menghadapi masalah tersebut, maka dari itu pada tugas akhir ini akan dibuat suatu sistem menggunakan metode *Naive Bayes* dan *URL Analysis* untuk mengklasifikasi sekumpulan *tweet* yang berisi *link* kedalam dua kelas yaitu *spam* dan *non spam*. Sebelum memasuki proses pengklasifikasian, sistem melakukan *preprocessing* terhadap dataset yang digunakan, hal tersebut bertujuan untuk meningkatkan performansi dari ke dua metode yang digunakan. Pada penelitian ini *Naive Bayes* dan *URL Analysis* menghasilkan nilai akurasi cukup tinggi yaitu 91% dan 93%.

Kata Kunci : *Twitter, Tweet, Spam, Naive Bayes, URL Analysis.*