

KLASIFIKASI E-MAIL SPAM MENGGUNAKAN METODE K-NEAREST NEIGHBOR

Alan Fitriyanto, Ilham Saifudin

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember
alanfitriyanto2@gmail.com, ilhamsaifudin@unmuhjember.ac.id

ABSTRAK

Email merupakan alat komunikasi penting yang digunakan oleh individu atau perusahaan saat ini. Adanya email yang mengandung spam sangat mengganggu bagi para pengguna email. Dalam penelitian ini peneliti melakukan klasifikasi email spam menggunakan algoritma K-NN guna mengetahui tingkat akurasi dan presisi dari algoritma ini dalam melakukan klasifikasi. Data dalam penelitian ini menggunakan dataset yang berasal dari *UCI Machine Learning* dengan total data 4601, terbagi atas 921 data validasi dan 3680 data uji serta latih. Dalam skenario uji penelitian ini menggunakan *K-Fold Cross Validation* dengan nilai k adalah 2, 4, 5, 8, 10. Hasil yang diperoleh dari penelitian ini adalah model alternatif yang didapat dari skenario uji dengan nilai akurasi tertinggi adalah pada uji dengan k=8 percobaan ke 8 yaitu 85.2% dan presisi 86.6%. sedangkan setelah diuji pada data validasi model ini menghasilkan tingkat akurasi 82.7% dan presisi 86.1%. hal ini menunjukkan bahwa terjadi penurunan tingkat akurasi sebesar 2.4% dan penurunan tingkat presisi sebesar 0.4%.

Kata kunci : Klasifikasi, *Email*, *Spam*, *K-NN*

