

ABSTRAK

DKI Jakarta merupakan sebuah kota di Indonesia yang memiliki padat penduduk yang tinggi yang harus kita perhatikan kondisi kesehatannya. Kualitas udara yang baik bisa menunjang produktivitas masyarakat DKI Jakarta untuk lebih aktif dan menciptakan kondisi yang segar. Banyak data yang muncul tentang kualitas udara pada DKI Jakarta yang selalu menurun karena berbagai faktor, oleh karena itu dibutuhkan sistem klasifikasi dengan penggunaan algoritma *naive bayes* yang dapat menunjang informasi kepada masyarakat setempat. Algoritma *naive bayes* adalah salah satu algoritma klasifikasi berdasarkan probabilitas yang membandingkan *data training* dan *data testing*. Keduanya dibandingkan melalui beberapa tahap persamaan, yang akhirnya diperoleh probabilitas tertinggi yang ditetapkan sebagai informasi. Membuat sebuah model uji pada klasifikasi menggunakan algoritma *naive bayes* yang bertujuan mencari hasil yang baik. Hasil dari pengklasifikasian pada data Indeks Standar Pencemaran Udara pada kota DKI Jakarta menghasilkan yaitu, dengan rata-rata akurasi 88%, *precision* 85%, *recall* 96%, *f1-score* 90%.

Kata Kunci: Klasifikasi, *Naive bayes*, *data mining*, Indeks Standar Pencemaran Udara, Udara