

Abstrak

Naïve Bayes merupakan salah satu teknik pengklasifikasian dalam data mining dengan menerapkan teorema Bayes dalam pengolahannya dan memberikan hasil optimal bila setiap atribut dalam dataset saling bebas. Namun pada umumnya, suatu *dataset* memiliki atribut numerik dan atribut nominal yang tidak saling bebas sehingga apabila dianggap saling bebas maka dapat menimbulkan permasalahan *classification error*. Oleh karena itu dibutuhkan suatu metode untuk meminimalkan nilai *error rate* pada kesalahan pengklasifikasian tersebut, salah satu metodenya adalah strategi pendiskritan. Diskritisasi adalah sebuah metode yang memetakan beberapa nilai numerik (X) ke sebuah interval bernilai nominal (X^*) berdasarkan pengaturan frekuensi di dalam satu interval sehingga bisa di dapatkan jumlah interval terbentuk dalam satu atribut numerik.

Salah satu metode pendiskritisasian yang diterapkan dalam Tugas Akhir ini adalah *Optimal Flexible Frequency Discretization* (OFFD) yang berbasis *sequential search* dan *wrapper based supervised* untuk *incremental learning*. Pada metode ini akan dilakukan *feature selection* dengan teknik *wrapper* untuk mendapatkan atribut yang optimal berdasarkan parameter *fmeasure* nya. Selanjutnya data dengan atribut optimal tersebut akan didiskrit secara *sequential search* untuk nilai frekuensi minimum tiap interval.

Kata kunci: *Wrapper based, Feature Selection, Diskritisasi, Sequential Search, Naïve Bayes Classifier, frekuensi interval, tingkat kesalahan*