

INTISARI

Supervised Classification Parsial dengan Algoritma S-EM (Spy Expectation Maximization) dan Metode Naive Bayes pada Dokumen Teks.

Oleh

TAN BRYAN LIANG

19/442584/PA/19333

Klasifikasi *Naive Bayes* merupakan salah satu metode dalam algoritma *machine learning*. Metode *Naive Bayes* menggunakan teorema Bayes untuk menentukan kelas dari setiap input. Dalam hal klasifikasi teks, metode *Naive Bayes* menentukan kelas setiap teks dengan diketahui sekumpulan dokumen teks beserta kelasnya. Dalam penelitian ini, diusulkan suatu teknik untuk menangani permasalahan *supervised classification* parsial. Dalam permasalahan tersebut, dokumen pelatihan hanya memberikan informasi mengenai satu dari dua kelas yang ada. Algoritma *S-EM (Spy Expectation Maximization)* diusulkan untuk mengatasi permasalahan *supervised classification* parsial tersebut. Algoritma *S-EM* mengirim sebagian dari sekumpulan dokumen pelatihan yang memiliki 1 kelas tersebut ke sekumpulan dokumen yang belum memiliki kelas untuk mendapatkan informasi mengenai kelas lainnya.

ABSTRACT

Partially Supervised Classification with *S-EM* (*Spy Expectation Maximization*)

Algorithm and *Naive Bayes* method of Text Documents

By

TAN BRYAN LIANG

19/442584/PA/19333

Naive Bayes classification is one of the methods in machine learning algorithms. The *Naive Bayes* method utilizes *Bayes* Theorem to determine the class of each input. In the case of text classification, the *Naive Bayes* method determines the class of each text given a set of text documents along with their classes. This study proposes a technique for handling partially supervised classification problems. In such problems, the training documents only provide information about one of the two existing classes. The *S-EM* (*Spy Expectation Maximization*) algorithm is proposed to address this issue. The *S-EM* algorithm sending a portion of the training document that has 1 class to the document that does not have a class to get information about the other class.