

**PENGARUH PEMBOBOTAN FITUR MENGGUNAKAN
INFORMATION GAIN TERHADAP PREDIKSI TINGKAT
KEGANASAN KANKER PAYUDARA**
(*THE EFFECT OF WEIGHTING FEATURES USING INFORMATION TO PREDICT
MALIGNANCY RATES OF BREAST CANCER*)

Muhammad Faris Faisal
Fakultas Teknologi Informasi dan Komunikasi
Universitas Semarang
faish3860@gmail.com

ABSTRACT

Anatomists will take samples of cancer cells that will be tested by the laboratory and calculated the value, but sometimes the more attributes that are calculated can reduce the level of accuracy in predicting the malignancy of breast cancer. To overcome this can be done with data mining with the feature weighting method, weighting the features themselves can produce the most influential attributes in making predictions, in this study we use Information Gain to weight features. Information Gain is one of the simplest feature weighting by calculating the entropy of each attribute so that it can be known the attribute value that will be recalculated. in this study we use Information Gain to weight features. Information Gain is one of the simplest feature weighting by calculating the entropy of each attribute so that it can be known the attribute value that will be recalculated. To implement this research, in this study the object used was the data of Wisconsin Breast Cancer (WBC) from the UCI This data repository came from Dr. William H. Woldberg (1989-1991), University of Wisconsin Hospitals Madison, USA. This data amounts to 699 with 11 attributes, the software that will be used is Rapid Miner. The purpose of this study was to determine the weighting attributes with Information Gain on the classification of cancer malignancies in the WBC data set.

Keywords: Feature Weighting, Information Gain, Breast Cancer

ABSTRAK

Kanker payudara merupakan salah satu penyebab kematian wanita terbesar didunia. Kanker payudara itu sendiri merupakan sel tumor yang terbentuk pada jaringan payudara. Salah satu upaya pencegahannya adalah dengan melakukan diagnosa terhadap kanker tersebut. Dalam melakukan diagnosa kanker payudara dokter ahli pantologi dan anatomi tubuh akan melakukan pengambilan sampel sel kanker yang akan di uji laboratorium dan dihitung nilai nya, namun terkadang semakin banyaknya atribut yang dihitung dapat mengurangi tingkat akurasi dalam melakukan prediksi tingkat keganasan kanker payudara. Untuk mengatasi hal ini dapat dilakukan dengan data minning dengan metode pembobotan fitur, pembobotan fitur sendiri dapat menghasilkan atribut yang paling berpengaruh dalam melakukan prediksi, dalam penelitian ini kami menggunakan *Information Gain* untuk melakukan pembobotan fitur. *Information Gain* merupakan salah satu pembobotan fitur paling sederhana dengan menghitung *entropy* total dikurangi *entropy* setiap atribut sehingga dapat diketahui nilai atribut yang paling berpengaruh. Untuk mengimplementasikan penelitian ini obyek yang digunakan adalah data *Wisconsin Breast Cancer* (WBC) dari UCI Repository data ini berasal dari Dr. William H. Woldberg (1989-1991), University of Winconsin Hospitals Madison, USA. Data ini berjumlah 699 dengan 11 atribut, *Software* yang akan digunakan adalah *Rapid Miner*. Tujuan penelitian ini untuk mengetahui pembobotan atribut dengan *Information Gain* terhadap prediksi tingkat keganasan kanker pada data set WBC.

Kata Kunci : Pembobotan Fitur, *Information Gain*, Kanker Payudara