

ABSTRACT

**THE COMPARISON OF CLASSIFICATION ACCURACY OF CERVICAL
CANCER CASES USING RANDOM FOREST AND SUPPORT
VECTOR MACHINE (SVM)
(CASE STUDY AT WISNUWARDHANA CANCER FOUNDATION
SURABAYA)**

Random Forest (RF) and Support Vector Machine (SVM) are computational-based statistical method. The former is used for classification cases, while the former is used for prediction and classification cases. SVM is a statistical method that can also identify cases of cervical cancer based on the results of examination by considering the factors of identification (risk factors). Meanwhile, RF is one of the methods used for classification by building many classification trees.

The general objective of this study was to compare RF and SVM performances in determining the classification of risk factors (age of first marriage, number of children, type of contraceptives, duration of contraceptive use, and patient age) affecting the incidence of cervical cancer.

The type of this research was unobstrusive research. The sample was the total population that had inclusion criteria, i.e. female patient who had performed papsmear at Wisnuwardhana Cancer Foundation Surabaya in January-December 2017.

SVM analysis results show that 1-APER was 0.995 or in other words, the right sample data were classified 99.5%. Sensitivity of SVM was 0.997 and the specificity was 0.988. Furthermore, the RF analysis results show that the calculation results of 1-APER was 1.00 or in other words, the right sample data were classified 100%. The sensitivity was 1.00 and the specificity was also 1.00.

The conclusion of this research is that all measures of accuracy that include 1-APER, sensitivity, specificity, and G-means in RF method is better than the support vector machine method. This indicates that RF method is better at classifying the cases of cervical cancer at Wisnuwardhana Cancer Foundation Surabaya.

Keywords : Support Vector Machine, Random Forest, papsmear, cervical cancer

ABSTRAK

**PERBANDINGAN KETEPATAN KLASIFIKASI KASUS KANKER
SERVIKS MENGGUNAKAN RANDOM FOREST DAN *SUPPORT
VECTOR MACHINE* (SVM)
(STUDI DI YAYASAN KANKER WISNUWARDHANA SURABAYA)**

Random Forest (RF) dan *Support Vector Machine* (SVM) adalah metode statistika berbasis komputasi. RF digunakan untuk kasus klasifikasi. Sedangkan SVM bisa digunakan untuk kasus prediksi dan klasifikasi. SVM merupakan metode statistik yang juga dapat mengidentifikasi kasus kanker serviks berdasarkan hasil pemeriksaan dengan memperhatikan faktor-faktor identifikasi (faktor risiko) yang melatarbelakangi. Sementara, RF merupakan salah satu metode yang digunakan untuk klasifikasi dengan membangun banyak pohon klasifikasi.

Tujuan umum penelitian ini adalah untuk membandingkan performansi RF dengan SVM dalam menentukan klasifikasi faktor risiko (umur pertama kali menikah, jumlah anak, jenis kontrasepsi, lama penggunaan kontrasepsi, dan umur pasien) yang berpengaruh terhadap kejadian kanker serviks.

Jenis penelitian *unobstrusive*. Sampel dalam penelitian ini merupakan total populasi yang memiliki kriteria inklusi, yaitu pasien wanita yang pernah melakukan *papsmear* di Yayasan Kanker Wisnuwardhana Surabaya dan pernah melakukan *papsmear* pada bulan Januari-Desember 2017.

Hasil analisis SVM menunjukkan bahwa 1-APER bernilai 0.995 atau dengan kata lain, data sampel yang tepat diklasifikasikan secara keseluruhan sebanyak 99.5%. *Sensitivity* bernilai 0.997 dan *specificity* bernilai 0.988. Selanjutnya, hasil analisis RF menunjukkan bahwa hasil perhitungan menunjukkan bahwa 1-APER bernilai 1.00 atau dengan kata lain, data sampel yang tepat diklasifikasikan secara keseluruhan sebanyak 100%. *Sensitivity* bernilai 1.00 dan *specificity* juga bernilai 1.00.

Kesimpulan dari penelitian ini adalah semua ukuran akurasi yang meliputi 1-APER, *sensitivitas*, *spesifisitas*, dan G-means pada metode *random forest* lebih baik dibandingkan dengan metode *support vector machine*. Hal ini menunjukkan bahwa metode *random forest* lebih bagus dalam melakukan pengklasifikasian pada kasus kanker serviks di Yayasan Kanker Wisnuwardhana Surabaya.

Kata Kunci : *Support Vector Machine*, *Random Forest*, *papsmear*, kanker serviks