

Aplikasi Diagnosis Penyakit Kanker Payudara Menggunakan Algoritma Sequential Minimal Optimization

Agung Wibowo*)

Program Studi Teknik Informatika, STMIK Nusa Mandiri Sukabumi
Jl. Veteran II No.20A, Kota Sukabumi, Indonesia 43114

Cara sitasi: A. Wibowo, "Aplikasi Diagnosis Penyakit Kanker Payudara Menggunakan Algoritma Sequential Minimal Optimazion," *Jurnal Teknologi dan Sistem Komputer*, vol. 5, no. 4, pp. 153-158, Okt. 2017. doi: 10.14710/jtsiskom.5.4.2017.153-158, [Online].

Abstract – Various methods for the diagnosis of breast cancer exist, but not many have been implemented as an application. This study aims to develop an application using SMO algorithm assisted by Weka to diagnose breast cancer. The application was web-based application and developed using Javascript. Test dataset and model formation used original Breast Cancer Database (WBCD) data without missing value. Test mode used 10-fold cross-validation. This application can diagnose breast cancer with an accuracy of 97.3645% and has a significant increase of accuracy than similar applications.

Keywords – Breast cancer; diagnosis algorithm; Weka; accuracy; SMO

Abstrak – Beragam metode untuk diagnosis kanker payudara, namun belum banyak yang diimplementasikan menjadi sebuah aplikasi. Penelitian ini bertujuan untuk mengembangkan aplikasi berdasarkan model hasil kalkulasi algoritma SMO berbantuan Weka untuk mendiagnosis penyakit kanker payudara. Aplikasi dikembangkan berbasis web menggunakan Javascript. Dataset pengujian dan pembentukan model menggunakan data Winconsin Breast Cancer Database original (WBCD) tanpa nilai hilang. Mode pengujian menggunakan 10-fold cross validation. Aplikasi ini dapat mendiagnosis kanker payudara dengan akurasi 97.3645% dan memiliki peningkatan akurasi yang signifikan dibandingkan aplikasi sejenis.

Kata Kunci – Kanker payudara; algoritma diagnosis; Weka; akurasi; SMO

I. PENDAHULUAN

Diagnosis penyakit menggunakan komputer memiliki banyak keuntungan. Diagnosis hasil pengolahan komputer dapat menjadi pendapat kedua bagi dokter dan berbiaya rendah. Selain itu, diagnosis berbasis komputer tersebut dapat menjadi solusi untuk meningkatkan akurasi diagnosis tanpa melalui prosedur medis [1].

Salah satu penyakit berbahaya adalah kanker. Bagi kaum perempuan, kanker yang paling menakutkan adalah kanker payudara. Perempuan yang teridentifikasi kanker payudara akan merasakan dampak secara psikologis, mulai dari keterkejutan mental, takut, depresi bahkan sampai panik dan bisa melakukan hal-hal yang tidak berarti dan sia-sia [2]. Kanker payudara ini merupakan penyakit *neoplasma* ganas dan berbahaya dengan jumlah penderita yang cukup tinggi, yaitu di urutan kedua terbanyak setelah kanker serviks di Indonesia [3] dan kanker paru di Dunia [4]. Penderita kanker payudara di stadium awal tidak merasakan gejala apa-apa [5]. Namun, kenyataannya pada tahun 2015 di Amerika terdapat lebih dari 226 ribu orang terdiagnosis sebagai penderita kanker payudara baru dan 39 ribu meninggal karenanya [6].

Penelitian tentang diagnosis kanker payudara berbantuan komputer telah banyak dilakukan. Aplikasi deteksi kanker payudara menggunakan algoritma fuzzy telah dikembangkan dalam [7]-[9]. Makhfudhoh [10] mengklasifikasikan kanker payudara menggunakan model jaringan syaraf tiruan fuzzy (JSN fuzzy). Peneliti lainnya masih berupa paparan analisa dan/atau komparasi akurasi diagnosis menggunakan algoritma regresi logistik [11], *Dempster Shafer* [11], *support vector machine* (SVM) [12], algoritma genetika [13], JSN [13], *certainty factor* [14], *Fuzzy MADM* [15], ID3 [16], C4.5 [3], *Naive Bayes* [3], [17], [18], *Sequential Minimal Optimization* (SMO) [3], dan *Multi Layer Preceptron* (MLP) [3]. Perbandingan metode diagnosis kanker payudara dan akurasinya ditunjukkan dalam Tabel 1. Sebagian metode tersebut diimplementasikan dalam bentuk aplikasi.

Data sumber penelitian yang digunakan dalam pengujian akurasi diagnosis kanker payudara tersebut sebagian besar menggunakan data publik dari UCI *Machine Learning Repository* dataset *breast cancer* [19]. Dataset kanker payudara di *repository* UCI terbagi menjadi 5, yaitu *breast cancer*, *breast Cancer Winconsin (original)* atau WBCD, *Breast Cancer Winconsin (Prognostic)* atau WPBC, *Breast cancer Winconsin (Diagnostic)* atau WDBC dan *Mammographic Mass*.

Safutra dan Prabowo [18] membuat aplikasi diagnosis kanker payudara berbasis desktop menggunakan algoritma *naive bayes classifier* dan

*) Penulis korespondensi (Agung Wibowo)
Email: agung_awo@nusamandiri.ac.id

Tabel 1. Penelitian diagnosis kanker payudara di Indonesia

Peneliti	Metode	Akurasi
Novianti dan Purnami [11]	Regresi logistik dan SVM	94,34%.
Zamani, Amaliah dan Munif [13]	JSN dan algoritma genetika	97%
Triaji [14]	<i>certainty factor</i> *	-
Primartha dan Fathiyah [7]	<i>fuzzy</i> metode Mamdani	91%
Zulkarnaini [15]	<i>fuzzy</i> MADM	90%
Mutlimah [8]	<i>fuzzy</i> *	85%
Makhfudhoh [10]	JSN <i>fuzzy</i> *	89%
Nabilah dan Kusumadewi [9]	<i>fuzzy inference system</i> *	-
Via, Nugroho dan Syafrizal [17]	<i>naive bayes</i>	97,82%
Wahyuni [3]	SMO, MLP, C4.5 dan <i>naive bayes</i>	97.6574%
Safutra dan Prabowo [18]	<i>naive bayes</i>	98,6726%
Dewa [12]	<i>Dempster Shafer</i>	86,56%

*Disertai aplikasi

dataset WBCD. Wahyuni [3] menguji uji akurasi diagnosis kanker payudara menggunakan algoritma SMO, MLP, C4.5 dan *Naive Bayes* pada dataset WBCD dan menunjukkan bahwa SMO memperoleh akurasi diagnosis tertinggi sebesar 97.6574%. Hasil uji akurasi SMO memiliki akurasi tertinggi didukung oleh Urmaliya dan Singhai [20] yang menyatakan bahwa akurasi diagnosis kanker payudara tertinggi adalah menggunakan algoritma SMO. Algoritma ini dapat dioptimasi dengan menggabungkan dengan algoritma seleksi fitur sehingga memperoleh akurasi 100%.

Penelitian mengenai kanker payudara di Indonesia cukup banyak seperti ditunjukkan pada Tabel 1. Namun, belum banyak yang menerapkan hasil penelitiannya menjadi sebuah aplikasi, seperti di antaranya dalam penelitian [8]-[10] dan [14]. Penelitian tersebut menggunakan algoritma *fuzzy* dan *certainty factor*. Penelitian ini bertujuan menerapkan model diagnosis algoritma SMO menjadi sebuah aplikasi diagnosis kanker payudara. Aplikasi ini diterapkan untuk mendeteksi kanker payudara dengan lebih akurat.

II. METODE PENELITIAN

Penelitian ini menggunakan algoritma SMO untuk mempercepat proses pelatihan dalam mendiagnosis penyakit kanker payudara [21]. Aplikasi ini dilakukan melalui lima tahap, yaitu pengambilan dataset pengujian

Tabel 2. Atribut WBCD dataset

Nama atribut	Domain	Tipe data
<i>Clump Thickness</i>	1 - 10	<i>Integer</i>
<i>Uniformity of Cell Size</i>	1 - 10	<i>Integer</i>
<i>Uniformity of Cell Shape</i>	1 - 10	<i>Integer</i>
<i>Marginal Adhesion</i>	1 - 10	<i>Integer</i>
<i>Single Epithelial Cell Size</i>	1 - 10	<i>Integer</i>
<i>Bare Nuclei</i>	1 - 10	<i>Integer</i>
<i>Bland Chromatin</i>	1 - 10	<i>Integer</i>
<i>Normal Nucleoli</i>	1 - 10	<i>Integer</i>
<i>Mitoses</i>	1 - 10	<i>Integer</i>

model, filter *missing value* dari dataset, uji model, perancangan aplikasi dan uji akurasi aplikasi.

Data yang digunakan dalam penelitian ini menggunakan dataset WBCD dari UCI *repository* [19]. Dataset WBCD memiliki atribut seperti yang terlihat pada Tabel 2. Distribusi kelasnya adalah *benign* (jinak) 458 atau 65,5% dan *malignant* (ganas) 241 atau 34,5% [22].

Dari filter 699 data, diperoleh data lengkap berjumlah 683 tanpa *missing value*. Rincian data filter adalah 458 jinak dengan 14 *missing value* dan 241 ganas dengan 2 *missing value*. Dataset hasil filter ini dijadikan data akhir yang digunakan untuk uji model dan berjumlah 683 data.

Data diuji menggunakan aplikasi *Waikato Environment for Knowledge Analysis* (WEKA) [23] dan dataset WEKA yang diunduh dari <http://repository.seasr.org/Datasets/UCI/arff/breast-w.arff>. Mode pengujian menggunakan *10-fold cross-validation* dengan membagi data mejadi 10 bagian dimana bagian pertama digunakan untuk pelatihan dan sembilan bagian lainnya digunakan untuk pengujian. Pengujian ini dilakukan di komputer dengan prosesor 1.66 Ghz dan memori 1 GB.

Model yang dihasilkan dari tahap sebelumnya diterjemahkan ke dalam kode program. Aplikasi dikembangkan berbasis web menggunakan bahasa *Javascript* agar bersifat multiplatform.

Pengujian akurasi dilakukan dengan cara membandingkan luaran diagnosis aplikasi dengan nilai hasil diagnosis dari setiap *record* pada dataset. Akurasi diagnosis aplikasi juga dibandingkan dengan hasil diagnosis model pada dataset WEKA.

Aplikasi yang dikembangkan dalam artikel ini menggunakan tiga aspek dari hasil uji model, yaitu model *machine learning*, simpulan, dan matrik galat bersama. Model *machine learning* digunakan untuk membentuk algoritma program. Simpulan digunakan sebagai dasar akurasi model dan bersama matrik galat digunakan sebagai uji akurasi aplikasi sehingga dapat dilihat apakah aplikasi yang dibangun memiliki peningkatan akurasi atau tidak.

Model *machine learning* dari hasil pengujian algoritma SMO akan memberikan bobot dari setiap atribut yang digunakan untuk klasifikasi. Bobot dari setiap atribut pada model diagnosis adalah seperti dinyatakan dalam Persamaan 1.

$$\begin{aligned}
& 1.5423 * (\text{normalized}) \text{ Clump_Thickness} \\
+ & 0.2231 * (\text{normalized}) \text{ Cell_Size_Uniformity} \\
+ & 1.317 * (\text{normalized}) \text{ Cell_Shape_Uniformity} \\
+ & 0.664 * (\text{normalized}) \text{ Marginal_Adhesion} \\
+ & 0.9158 * (\text{normalized}) \text{ Single_Epi_Cell_Size} \\
+ & 1.5121 * (\text{normalized}) \text{ Bare_Nuclei} \\
+ & 1.2144 * (\text{normalized}) \text{ Bland_Chromatin} \\
+ & 0.7219 * (\text{normalized}) \text{ Normal_Nucleoli} \\
+ & 1.1801 * (\text{normalized}) \text{ Mitoses} \\
- & 2.5757
\end{aligned} \tag{1}$$

Nilai normalisasi didapatkan dengan menggunakan fungsi normalisasi yang digunakan Weka [23]. Fungsi normalisasi data yang digunakan oleh Weka adalah *min-max* [24]. Fungsi ini menormalisasi semua nilai numerik dalam dataset yang diberikan terlepas dari atribut kelas. Nilai yang dihasilkan untuk data yang digunakan saat menghitung interval normalisasi secara *default* adalah [0,1]. Nilai atribut [a,b,c,...,i] dinyatakan dengan *x*. Parameter *min* setiap atribut adalah 0 dan *max* bernilai 10. Nilai *scale* default adalah 1, sedangkan *translation* bernilai default 0. Fungsi normalisasinya menjadi seperti yang dinyatakan dalam Persamaan 2.

$$\begin{aligned}
F(x) &= \left(\frac{x - \min(x)}{\max(x) - \min(x)} \right) * \text{scale} + \text{translation} \\
&= \left(\frac{x - 1}{10 - 1} \right) * 1 + 0
\end{aligned} \tag{2}$$

III. HASIL DAN PEMBAHASAN

Hasil dari uji data dengan menggunakan aplikasi Weka dan 10-fold *cross validation* ditunjukkan dalam Tabel 3. Akurasi algoritma SMO yang diperoleh adalah sebesar 97.0717 % dan rerata waktu olah adalah 1,04 detik (prosesor 1,66 GHz, RAM 1 GB). *Kappa Statistic* menyatakan perbandingan akurasi teramati dengan akurasi yang diharapkan (peluang acak). *Mean absolute error* menunjukkan rerata absolut galat antara nilai yang diamati dengan yang diprediksi. *Root mean squared error* mengukur perbedaan antara nilai (sampel dan nilai populasi) yang diprediksi oleh model atau estimator dengan nilai yang benar-benar diamati. *Relative absolute error* menyatakan rasio absolut galat pengukuran terhadap nilai pengukuran yang diterima. *Root relative absolute error* dihitung dengan membagi RMSE oleh RMSE yang diperoleh dengan hanya meramalkan rata-rata nilai target (dan kemudian mengalikan dengan 100). *Root relative squared error* dihitung sebagai MEA dibagi dengan kesalahan classifier Zero.

Akurasi hasil pengujian menggunakan Weka akan dijadikan pembandingan dengan akurasi aplikasi. Setelah model didapatkan, selanjutnya adalah menghitung nilai normalisasi dari setiap *record* dalam dataset. Hasil perhitungan normalisasi didapatkan data perbandingan nilai maksimum, minimum, rerata dan kuartil dari data kanker jinak dan ganas. Tabel 4 menunjukkan data statistik sebaran nilai hasil normalisasi dan Gambar 1 menunjukkan sebarannya dalam bentuk grafik.

Tabel 3. Hasil pengujian statistik model Weka

	<i>Stratified cross-validation</i>	<i>Summary</i>
<i>Correctly Classified Instances</i>	663	97.0717%
<i>Incorrectly Classified Instances</i>	20	2.9283 %
<i>Kappa statistic</i>	0.9359	
<i>Mean absolute error</i>	0.0293	
<i>Root mean squared error</i>	0.1711	
<i>Relative absolute error</i>	6.4342 %	
<i>Root relative squared error</i>	35.8782 %	
<i>Total Number of Instances</i>	683	

Tabel 4. Statistik normalisasi data kanker jinak dan ganas

	<i>Jinak</i>	<i>Ganas</i>
<i>MIN</i>	-2.5757	-1.10707
<i>MAX</i>	2.247878	5.987111
<i>Average</i>	-1.80644	2.604721
<i>Q1</i>	-2.20408	1.775272
<i>Q2</i>	-1.89452	2.600172
<i>Q3</i>	-1.5987	3.393222

Tabel 5. Data akurasi diagnosis aplikasi

<i>Diagnosis</i>	<i>Simpulan</i>
Jinak (salah diagnosis)	17
Ganas (salah diagnosis)	1
Akurasi (Benar)	97.3645680819912
Rerata Selisih Galat	0.206912037037037

Nilai batas atau vonis diagnosis antara jinak dan ganas pada aplikasi ini dihitung berdasarkan hasil bagi antara Q3 Jinak dan Q1 ganas. Nilai vonisnya adalah -0.711063888888889. Pengujian akurasi aplikasi hitung berdasarkan hasil analisa menggunakan nilai batas vonis dan didapatkan data akurasi diagnosis aplikasi, yang dapat dilihat pada Tabel 5. Nilai vonis, fungsi normalisasi (Persamaan 2) dan model algoritma SMO (Persamaan 1) digunakan dalam aplikasi seperti dinyatakan dalam Program 1.

Program 1. Algoritma aplikasi

```

ALGORITMA Diagnosis_kanker_payudara_SMO
{IS : diberikan 9 data atribut kanker payudara
  untuk di diagnosis dengan selang nilai 1..10}
{FS : ditampilkan data hasil diagnosis bds model
  algoritma SMO}

```

```

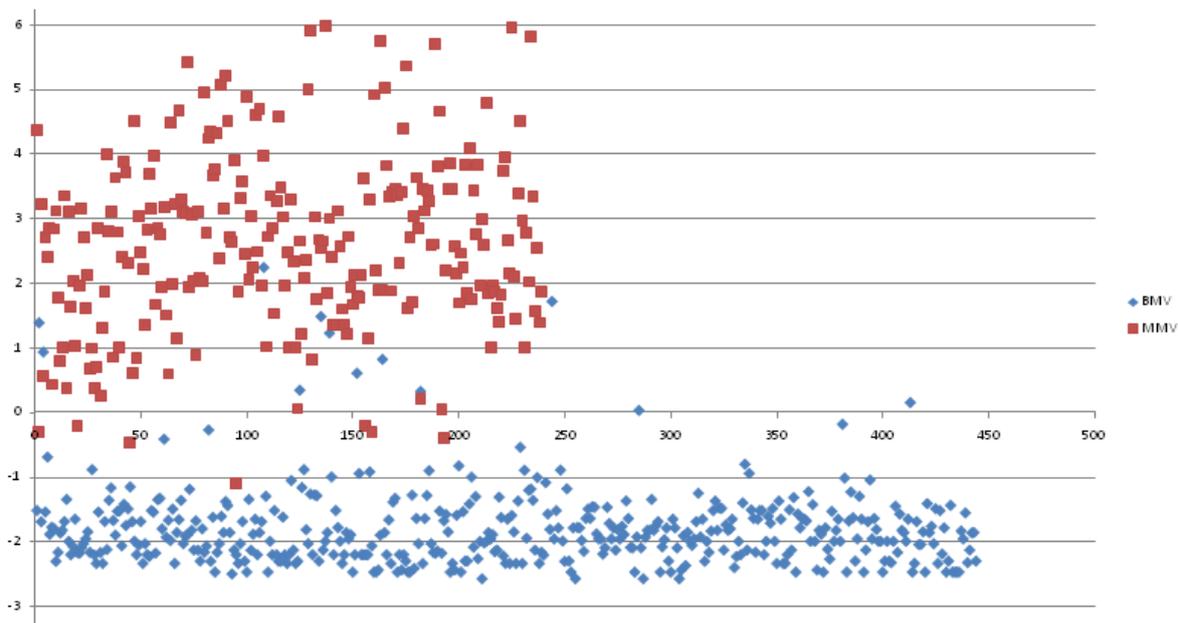
KAMUS DATA
hitung : real;
a, b, c, d, e, f, g, h, i: integer; //masukan nilai
  atribut diagnosis

```

```

FUNCTION normalized(x)
BEGIN
  return ((x-1)/(10-1))*1+0;
END

```



Gambar 1. Sebaran Normalisasi Data (biru = jinak, merah = ganas)

```
//PROGRAM
INPUT (a); //Clump_Thickness
INPUT (b); //Cell_Size_Uniformity
INPUT (c); //Cell_Shape_Uniformity
INPUT (d); //Marginal_Adhesion
INPUT (e); //Single_Epi_Cell_Size
INPUT (f); //Bare_Nuclei
INPUT (g); //Bland_Chromatin
INPUT (h); //Normal_Nucleoli
INPUT (i); //Mitoses

//model algoritma SMO
hitung = (1.5423*normalized(a))
+ (0.2231*normalized(b))
+ (1.317 *normalized(c))
+ (0.664 *normalized(d))
+ (0.9158 *normalized(e))
+ (1.5121 *normalized(f))
+ (1.2144 *normalized(g))
+ (0.7219 *normalized(h))
+ (1.1801 *normalized(i))
- 2.5757;

//Diagnosis
IF ( hitung >= -0.711063888888889) THEN
  OUTPUT( " Malignant");
ELSE OUTPUT(" Benign")
```

Tampilan aplikasi diagnosis kanker payudara dan hasil kalkulasi dengan contoh masukan pengguna dapat dilihat pada Gambar 2. Tampilan tersebut merupakan tampilan dasar dan dapat disesuaikan dengan kebutuhan. Kode program lengkap dari aplikasi ini tersedia dan dapat diunduh bebas pada laman <http://jsbin.com/vamubaqide/edit?html,output>.

Nilai akurasi diagnosis antara hasil pengujian Weka dan hasil diagnosis aplikasi dapat dilihat pada Tabel 6. Aplikasi diagnosis kanker payudara menggunakan algoritma SMO dengan nilai vonis hasil dari nilai tengah Q3 jinak dan Q1 ganas mengalami peningkatan akurasi dari yang ditunjukkan Weka sebesar 0,29287%.



Gambar 2. Aplikasi diagnosis dan hasil kalkulasi

Tabel 6. Komparasi akurasi Weka dan aplikasi

	<i>Weka</i>		<i>Aplikasi</i>	
	Data	Simpulan	Data	Simpulan
<i>Correctly Classified Instances</i>	663	97.0717%	665	97.36457%
<i>Incorrectly Classified Instances</i>	20	2.9283 %	18	2.635432%

Tabel 7. Matrik galat

WEKA		Aplikasi		Hasil Diagnosis
a	b	a	b	
432	12	427	17	a = jinak
8	231	1	238	b = ganas

Uji akurasi berikutnya adalah dengan menggunakan matrik galat. Matrik galat digunakan untuk menganalisis seberapa baik algoritma klasifikasi mengenali tupel dari kelas yang berbeda [25]. Tabel 7

menunjukkan perbandingan matriks galat diagnosis kanker menurut hasil olah Weka dan aplikasi yang dihasilkan. Nilai matriks galat menurut Weka ada 12 data jinak diklasifikasikan sebagai ganas dan ada 8 data ganas yang diklasifikasikan sebagai jinak. Hasil uji diagnosis aplikasi diperoleh sebanyak 17 data jinak yang didiagnosis sebagai ganas dan ada 1 data kanker ganas didiagnosis jinak. Perbaikan perhitungan nilai vonis ideal untuk kalkulasi diagnosis masih perlu dilakukan untuk mengurangi kesalahan diagnosis kanker jinak yang masih besar.

Dari hasil dalam Tabel 6 dan Tabel 7, terlihat bahwa aplikasi ini dapat menghasilkan peningkatan deteksi kanker ganas lebih baik dibandingkan model hasil uji data menggunakan Weka [23]. Dibandingkan dengan akurasi metode diagnostik dalam Tabel 1, aplikasi memiliki akurasi lebih rendah dibandingkan dengan hasil penelitian Wahyuni [3] yang memperoleh akurasi tertinggi dengan SMO, dan dengan penelitian Via dkk. [9] serta Safutra dan Prabowo [18] yang menggunakan metode *naive bayes*, namun lebih akurat daripada [7], [8], [10]-[13] dan [15] yang menggunakan fuzzy, regresi logistik dan SVM, serta Dempster Shafer. Untuk mendapatkan akurasi yang lebih baik lagi, seleksi fitur dapat ditambahkan dalam aplikasi seperti dalam [20]. Dihasilkannya aplikasi diagnosis kanker payudara yang akurat dan cepat dengan waktu rerata 1,04 detik ini dapat membantu dokter atau analis kesehatan untuk mendiagnosis kanker payudara dengan lebih cepat.

IV. KESIMPULAN

Penelitian ini telah menghasilkan aplikasi diagnosis kanker payudara menggunakan SMO yang lebih akurat dibandingkan model uji Weka. Aplikasi ini juga menghasilkan diagnosis yang lebih akurat dari aplikasi sejenis yang menggunakan metode fuzzy.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Dr. William H. Wolberg, Olvi Mangasarian selaku kreator dan donatur dataset yang digunakan dalam penelitian ini.

DAFTAR PUSTAKA

- [1] A. O. Bilska-Wolak, C. E. Floyd Jr., L. W. Nolte, and J. Y. Lo, "Application of Likelihood Ratio to Classification of Mammographic Masses Performance Comparison to Case-based Reasoning," *Medical Physics*, vol. 10, no. 5, pp. 949-958, Mei 2003.
- [2] A. F. Rahmah and E. L. Widuri, "Post Traumatic Growth pada Penderita Kanker Payudara," *Humanitas*, vol. 8, no. 2, pp. 114-128, Agustus 2011.
- [3] E. S. Wahyuni, "Penerapan Metode Seleksi Fitur untuk Meningkatkan Hasil Diagnosis Kanker Payudara," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 7, no. 1, pp. 283-294, April 2016.
- [4] S. Gupta, P. F. Chyn, and M. K. Markey, "Breast Cancer CAD x Based on BI-RADS™ Descriptors from Two Mammographic Views," *Medical Physics*, vol. 33, no. 6, pp. 1810-1817, Juni 2006.
- [5] M. Elter, R. Schulz-Wendtland, and T. Wittenberg, "The Prediction of Breast Cancer Biopsy Outcomes using Two CAD Approaches that Both Emphasize an Intelligible Decision Process," *Medical Physics*, vol. 34, no. 11, pp. 4164-4172, November 2007.
- [6] H. Cho et al., "A Similarity Study of Content-based Image Retrieval System for Breast Cancer using Decision Tree," *Medical Physics*, vol. 40, no. 1, pp. 012901-1 - 012901-13, January 2013.
- [7] R. Primartha and N. Fathiyah, "Sistem Pakar Fuzzy Untuk Diagnosis Kanker Payudara Menggunakan Metode Mamdani," *Generic*, vol. 8, no. 1, pp. 190-197, Maret 2013.
- [8] M. Mutlimah, "Aplikasi Sistem Fuzzy untuk Diagnosa Kanker Payudara (Breast Cancer)," Universitas Negeri Yogyakarta, Yogyakarta, Desertasi 2014.
- [9] G. P. Nabilah and S. Kusumadewi, "Fuzzy Inference System untuk Penentuan Resiko Kanker Payudara," in *SNATI*, Kudus, 2015, pp. 101-109.
- [10] N. U. Makhfudhoh, "Klasifikasi Kanker Payudara dari Citra Mammografi Menggunakan Model Fuzzy Neural Network," Disertasi, Universitas Negeri Yogyakarta, Yogyakarta, 2014.
- [11] F. A. Novianti and S. W. Purnami, "Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi Fourina Ayu Novianti dan Santi Wulan Purnami," *Jurnal Sains dan Seni*, vol. 1, no. 2, pp. D-147 - D-152, September 2012.
- [12] N. K. Dewa, "Rancang Bangun Ssistem Pakar Diagnosis Penyakit Kanker Payudara Menggunakan DEMPSTER - SHAFER," Disertasi, Universitas Pesantren Tinggi Darul Ulum, Jombang, 2016.
- [13] A. M. Zamani, B. Amaliah, and A. Munif, "Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara," *Jurnal Teknik ITS*, vol. 1, no. 1, pp. A-222 - A-227, September 2013.
- [14] M. Triaji, "Sistem Pakar Pendeteksi Kanker Payudara Menggunakan Metode Certainty Factor," Skripsi, Universitas Dian Nuswantoro, Semarang, 2012.
- [15] I. Zulkarnaini, "Diagnosa Penyakit Kanker Payudara dengan Menggunakan Sistem Fuzzy Multi Attribute Decision Making," *LENTERA*, vol. 13, no. 2, pp. 53-56, Juni 2013.
- [16] P. A. Maulida, "Decision Tree dengan Algoritma ID3 untuk Melakukan Deteksi," Skripsi, Universitas Dian Nuswantoro, Semarang, 2015.
- [17] Y. V. Via, B. Nugroho, and A. Syafrizal, "Sistem Pendukung Keputusan Klasifikasi Tingkat Keganasan Kanker Payudara dengan Metode

- Naive Bayes Classifier," *SCAN*, vol. 10, no. 2, pp. 63-68, Januari 2015.
- [18] A. R. Safutra and D. W. Prabowo, "Diagnosis Penyakit Kanker Payudara Menggunakan Metode Naive Bayes Berbasis Desktop," *Jurnal Penelitian Dosen FIKOM (UNDA)*, vol. 6, no. 1, pp. 1-6, Oktober 2016.
- [19] Center for Machine Learning and Intelligent Systems, "UCI Machine Learning Repository," [Online]. Available: <https://archive.ics.uci.edu/ml/datasets.htm>. [Diakses: Okt. 26, 2017].
- [20] A. Urmaliya and J. Singhai, "Sequential Minimal Optimization for Support Vector Machine with Feature Selection in Breast Cancer Diagnosis," in *Proc. 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, Shantala, India, 2013, pp. 481-486.
- [21] J. C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines," Microsoft Research, Technical Report MSR-TR-98-14, 1998.
- [22] O. L. Mangasarian and W. H. Wolberg, "Cancer Diagnosis via Linear Programming," *SIAM News*, vol. 23, no. 5, pp. 1-18, September 1990.
- [23] WEKA. (2010, Maret) SourceForge. [Online]. <http://weka.sourceforge.net/doc.stable/weka/filters/unsupervised/attribute/Normalize.html>
- [24] Hitachi Group Company. (2008, Desember) Pentaho. [Online]. <http://wiki.pentaho.com/display/DATAMINING/Normalize+%28attribute%29>
- [25] J. Han, M. Kamber and J. Pei, *Data Mining Concepts and Techniques, 3rd ed.*, Morgan Kaufman, Ed. Waltham, Massachusetts: Elsevier Inc, 2011.