

Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine

Chalifa Chazar¹, Bagus Erawan Widhiaputra²

Program Studi Teknik Informatika^{1,2}

STMIK Indonesia Mandiri, Jl. Jakarta No.79 Bandung^{1,2}

Email: chalifa.chazar@gmail.com¹, bwidhiaputra@gmail.com²

ABSTRAK

Kanker payudara merupakan penyebab kematian nomor dua pada wanita. Penyakit ini sulit dideteksi pada fase awal. Akan tetapi, kebanyakan penderita baru mengetahui kondisinya setelah memasuki fase tertentu dalam kondisi yang parah dan sulit disembuhkan. Salah satu bentuk pemeriksaan untuk mendiagnosis penyakit kanker payudara adalah dengan melakukan biopsi. Biopsi adalah teknik pemeriksaan yang dilakukan dengan mengambil cairan di payudara menggunakan Fine Needle Aspiration (FNA), selanjutnya hasil biopsi FNA akan diperiksa lagi di laboratorium untuk mendapatkan hasil diagnosis. Untuk mendapatkan hasil yang akurat dari proses biopsi dibutuhkan waktu yang lama. Machine Learning (ML) dapat digunakan untuk mencari dan menemukan pola yang unik dari sekumpulan data. Algoritma Support Vector Machine (SVM) dipilih karena algoritma ini mampu mengklasifikasikan nilai ke dalam kelas-kelas tertentu. Algoritma SVM juga memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan algoritma lainnya. Penelitian ini bertujuan untuk membangun aplikasi ML yang dapat mendiagnosis penyakit kanker payudara dengan menggunakan Algoritma SVM untuk mencari pola data dari sekumpulan data masa lalu untuk menghasilkan hasil diagnosis yang akurat. Hasil dari penelitian ini menunjukkan bahwa Algoritma SVM pada ML dapat digunakan untuk mencari suatu pola data dari sekumpulan data masa lalu yang dapat menghasilkan prediksi untuk menentukan sel hidup kanker payudara bersifat ganas atau jinak.

Kata Kunci: Kanker payudara, Machine Learning, Support Vector Machine

ABSTRACT

Breast cancer is the second leading cause of death in women. This disease is difficult to detect in the initial phase. However, most sufferers only know their condition after entering a certain phase in a condition that is severe and difficult to cure. One form of examination to diagnose breast cancer is to do a biopsy. Biopsy is an examination technique performed by taking fluid in the breast using Fine Needle Aspiration (FNA), then the results of the FNA biopsy will be examined again in the laboratory to get the diagnosis. To get accurate results from the biopsy process requires a long time. Machine Learning (ML) can be used to search for and find unique patterns from a set of data. Support Vector Machine (SVM) algorithm was chosen because this algorithm is able to classify values into certain classes. SVM algorithm also has a higher level of accuracy compared to other algorithms. This research aims to build an ML application that can diagnose breast cancer by using the SVM Algorithm to look for data patterns from a collection of past data to produce an accurate diagnosis. The

results of this research indicate that the SVM Algorithm in ML can be used to search for a pattern of data from a set of past data that can produce predictions for determining which breast cancer cells are malignant or benign.

Keyword: *Breast Cancer, Machine Learning, Support Vector Machine*

1. PENDAHULUAN

Machine Learning is a field of komputer science that gives computer the ability to learn without being explicitly programmed (Primartha, 2018:12). *Machine Learning* (ML) adalah salah satu aplikasi dari *Artificial Intilligent* (AI) yang fokus kepada pengembangan sebuah sistem yang mampu belajar sendiri tanpa harus diprogram berulang kali. ML membutuhkan sebuah data (*data traning*) sebagai proses *learning* sebelum menghasilkan sebuah hasil. Jadi, secara sederhana dapat dijelaskan bahwa *Machine Learning* adalah pemograman komputer untuk mencapai kriteria/performa tertentu dengan menggunakan sekumpulan data *traning* atau pengalaman di masa lalu (*past experience*) (Primartha, 2018:13). Beberapa penelitian yang telah dilakukan, menyimpulkan ML dapat digunakan pada bidang medis untuk memprediksi penyakit (Zhao et al., 2006)(Son, Kim, Kim, Choi, & Lee, 2010)(Janardhanan, Heena, & Sabika, 2015)(Asri, Mousannif, Al Moatassime, & Noel, 2016).

Hingga saat ini terdapat beberapa algoritma ML yang dapat digunakan dan dikembangkan untuk berbagai tujuan. Sebuah penelitian dilakukan untuk membandingkan efektifitas beberapa algoritma dalam ML antara lain *Naïve Bayes* (NB), *Radial Basis Function* (RBF) dan *Support Vektor Machine* (SVM) (Janardhanan et al., 2015). Berdasarkan hasil penelitian, diperoleh bahwa SVM adalah algoritma yang paling tinggi tingkat akurasi. Dibuktikan oleh tes algoritma tersebut yang menunjukkan nilai akurasi algoritma SVM mencapai 93,75% sedangkan untuk algoritma NB hanya 71,67% dan RBF sebesar 70,01% (Janardhanan et al., 2015).

The support vector machine (SVM) is a relatively new classification or prediction method developed by Cortes and Vapnik in the 1990s as a result of the collaboration between the statistical and the machine learning research community (Son et al., 2010). SVM adalah suatu teknik klasifikasi untuk masalah nonlinier. *SVM is most suitable for working accurately and efficiently with high*

dimensionality feature spaces in addition to that SVM is based on strong mathematical foundations and results in simple way and very powerful algorithms (Edriss, Ali, & Feng, 2016).

Kanker payudara atau *Carsinoma Mammae* adalah kondisi ketika sel kanker terbentuk di jaringan payudara. Kanker bisa terbentuk di kelenjar yang menghasilkan susu (*lobulus*), atau di saluran (*duktus*) yang membawa air susu dari kelenjar ke puting payudara, kanker juga bisa terbentuk di jaringan lemak atau jaringan ikat di dalam payudara. Kanker payudara merupakan penyebab utama tingkat kematian kedua pada wanita. *Breast cancer represents about 12% of all new cancer cases and 25% of all cancers in women* (Asri et al., 2016). Organisasi kesehatan dunia, *International Agency for Research on Cancer* (IARC) memperkirakan lebih dari 400.000 wanita meninggal setiap tahun karena kanker payudara (Edriss et al., 2016). Masih banyak orang yang berasumsi bahwa kanker sama seperti tumor, kenyataannya tumor yang muncul tidak selalu menjadi kanker. Pemeriksaan dengan biopsi dan *mammografi* dapat digunakan untuk mendeteksi jenis kanker payudara. Hasil pemeriksaan biopsi dengan *Fine Needle Aspiration* (FNA) dapat menentukan jenis sel kanker payudara apakah bersifat ganas atau jinak. Pertumbuhan kanker payudara yang berawal dari tumor dikelompokkan menjadi beberapa stadium, mulai dari stadium 0 sampai IV. Keterlambatan dalam mendeteksi gejala kanker payudara menyebabkan banyak penderita baru mengetahui kondisinya setelah memasuki stadium yang tinggi (rata-rata pada stadium III dan IV). Pada kondisi ini, risiko kematian menjadi jauh lebih tinggi.

Penelitian ini bertujuan untuk membangun sebuah aplikasi ML dengan menggunakan algoritma SVM yang dapat digunakan untuk mendiagnosis penyakit kanker payudara berdasarkan pola data yang diambil dari biopsi. Hasil diagnosis dapat menghasilkan suatu prediksi penentuan jenis sel kanker payudara bersifat ganas atau jinak.

2. METODE PENELITIAN

1. Diagnosis Kanker payudara

Kanker payudara terbentuk ketika sel-sel di dalam payudara tumbuh secara abnormal dan tidak terkendali. Sel tersebut umumnya membentuk tumor. Kanker

payudara dapat dideteksi dengan melakukan pemeriksaan biopsi dan *mammografi*. (1) Biopsi adalah teknik pemeriksaan yang dilakukan dengan mengambil cairan pada payudara dengan menggunakan FNA, kemudian hasil biopsi FNA akan diperiksa kembali di laboratorium untuk mendapatkan hasil diagnosis. (2) *Mammografi* adalah teknik pemeriksaan dengan menggunakan Sinar X-ray dengan kadar rendah untuk menilai jaringan payudara. Biopsi merupakan pemeriksaan yang paling akurat untuk mendeteksi kanker payudara, karena mampu memperlihatkan jenis sel kanker dan tingkat stadiumnya. Kanker payudara dapat dinilai dengan tiga faktor utama yaitu tumor (*primary tumor*), kelenjar getah bening (*regional lymph nodes*), dan penyebaran (*distant metastasis*) (AJCC, 2017). Tabel 1, menjelaskan indikator diagnosis kanker payudara.

Tabel 1. Indikator diagnosis kanker payudara

Primary Tumor (T)	
T0	No evidence of primary tumor
Tis	Carcinoma in situ
T1, T2, T3, T4	Increasing size and/or local extension of the primary tumor
TX	Primary tumor cannot be assessed (use of TX should be minimized)
Regional Lymph Nodes (N)	
N0	No regional lymph node metastases
N1, N2, N3	Increasing number or extend of regional lymph node involvement
NX	Regional lymph nodes cannot be assessed (use of NX should be minimized)
Distant Metastasis (M)	
M0	No distant metastases
M1	Distant metastases present

Sumber: American Join Committee on Cancer (2010)

Hasil biopsi selanjutnya dibawa ke laboratorium untuk pemeriksaan lebih detail. FNA akan menunjukkan beberapa indikator yang berhubungan dengan perkembangan sel kanker payudara. Tabel 2, menjelaskan deksripsi indikator dari hasil FNA.

Tabel 2. Deskripsi indikator hasil FNA pada Breast Cancer Wisconsin

No	Indikator	Deskripsi
1	Clump Thickness	Menilai apakah sel tersebut mono atau multi-layer
2	Uniformity of cell size	Evaluasi konsistensi ukuran sel dalam sample
3	Uniformity of cell shape	Evaluasi konsistensi bentuk sel dalam sample
4	Marginal Adhesion	Menghitung proporsi sel yang saling menyatu
5	Single Epithelial Cell Size	Mengukur pembesaran ukuran sel epitel
6	Bare Nuclei	Proporsi nuclei yang dikelilingi oleh sitoplasma melawan yang tidak
7	Bland Chromatin	Menilai samanya “tekstur” dari nucleus dalam rentang halus hingga kasar
8	Normal Nucleoli	Menentukan apakah nucleoli kecil dan hampir tidak terlihat atau lebih terlihat
9	Mitoses	Menggambarkan tingkat aktivitas mitoses

2. Dataset Breast Cancer Wisconsin

Data yang digunakan pada penelitian ini menggunakan *dataset Breast Cancer Wisconsin (Diagnostic)*, yang dibuat oleh Dr. William H. Wolberg, W. Nick Street, dan Olvi L. Mangasarian. Database tersebut berupa *dataset* yang didapat dari hasil analisis citra digital massa payudara dengan menggunakan metode FNA, mereka menganalisis perkembangan sel hidup yang abnormal pada citra digital.

Tabel 3. Informasi dataset Breast Cancer Wisconsin

#	Attribute	Domain
H1	Sample code number	id number
H2	Clump Thickness	1 -10
H3	Uniformity of Cell Size	1 -10
H4	Uniformity of Cell Shape	1 -10
H5	Marginal Adhesion	1 -10
H6	Single Epithelial Cell Size	1 -10
H7	Bare Nuclei	1 -10
H8	Bland Chromatin	1 -10
H9	Normal Nucleoli	1 -10
H10	Mitoses	1 -10
H11	Class:	(2 for benign, 4 for malignant)

Sumber: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>

Tabel 3, menjelaskan tentang informasi dari *dataset Breast Cancer Wisconsin*. *Breast Cancer Wisconsin* memiliki 699 *instance* (*benign*: 458 dan *malignant*: 241), 2 kelas (65.5% ganas dan 34.5% jinak), dan 11 atribut bernilai *integer* (Asri et al., 2016). *Dataset* ini juga merupakan indikator yang dapat dilihat pada sel hidup untuk mendeteksi adanya kanker payudara, setiap *record* memiliki sembilan atribut indikator selain *Sample Code Number* dan *Class* yang sembilan atribut tersebut dinilai pada skala interval 1 sampai 10, dengan skala 10 merupakan penilaian keadaan paling abnormal, sehingga semakin dekat nilai dari masing-masing atribut sampai ke 10 maka semakin besar kemungkinan terdeteksi *malignant* (ganas). Tabel 4, menggambarkan *sample dataset* dari *Breast Cancer Wisconsin*.

Tabel 4. Contoh dataset Breast Cancer Wisconsin

H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11
1000025	5	1	1	1	2	1	3	1	1	12
1002945	5	4	4	5	7	10	3	2	1	2
1015425	3	1	1	1	2	2	3	1	1	2
1016277	6	8	8	1	3	4	3	7	1	2
1017023	4	1	1	3	2	1	3	1	1	2
1017122	8	10	10	8	7	10	9	7	1	4
1018099	1	1	1	1	2	10	3	1	1	2
1018561	2	1	2	1	2	1	3	1	1	2
1033078	2	1	1	1	2	1	1	1	5	2
1033078	4	2	1	1	2	1	2	1	1	2

Sumber: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>

3. Algoritma SVM

Konsep SVM bertujuan untuk menemukan *hyperplane* yang memisahkan himpunan data ke dalam dua kelas secara linier (Suyanto, 2018:99). *Hyperplan* adalah istilah yang dibuat general untuk semua dimensi. SVM berusaha untuk menemukan *hyperplane* yang paling optimum atau terbaik. Proses klasifikasi memiliki dua proses, yaitu (1) proses *training* dan (2) proses *testing*. Proses *training* digunakan untuk membangun model dari suatu *training* set. Algoritma SVM pada proses *training* dilakukan dengan delapan langkah penyelesaian. Berikut ini adalah langkah-langkah yang dilakukan.

(1) Inisiasi nilai

$$\begin{aligned}\alpha &= 0 \\ C &= 1 \\ \varepsilon &= 0,001 \\ \gamma &= 0,5 \\ \lambda &= 0,5\end{aligned}$$

(2) Hitung *Dot Product* data *training*

$$A x A^T$$

(3) Hitung matriks

$$D_{ij} = y_i y_j (K(\vec{x}_i, \vec{x}_j) + \lambda^2)$$

(4) Cari nilai E (Error)

$$E_i = \sum_{j=1}^l \alpha_j D_{ij}$$

(5) Hitung $\Delta\alpha$

$$\Delta\alpha_i = \min \{ \max[\gamma(1 - E_1), -\alpha_j], C - \alpha_i \}$$

(6) Jika $\Delta\alpha < \varepsilon$ maka iterasi selanjutnya

$$\Delta\alpha = \alpha_i + \Delta\alpha$$

(7) Mencari nilai w^+ dan w^-

$$w^+ = \Delta\alpha \cdot D_{ij^+}$$

$$w^- = \Delta\alpha \cdot D_{ij^-}$$

(8) Mencari nilai bias (b)

$$b = -\frac{1}{2} ([\vec{w} \cdot \vec{x}_{-1}] + [\vec{w} \cdot \vec{x}_{+1}])$$

Proses *testing* digunakan untuk menguji keakuratan model yang dihasilkan dari proses *training*. Algoritma SVM pada proses *testing* dilakukan dengan dua langkah penyelesaian. Berikut ini adalah langkah-langkah yang dilakukan.

(1) Hitung *Dot Product* data *sample* dengan semua data *training*

$$K(x_i, y) = x \cdot y$$

(2) Hitung keputusan, jika nilai positif maka masuk kelas 1, jika negatif maka masuk kelas -1

$$F(x) = \sum_{i=1}^m (\alpha_i \cdot y_i \cdot K(x, x_i)) + b$$

3. HASIL DAN PEMBAHASAN

1. Proses Training

Proses biopsi membutuhkan waktu yang lama, sebelum menghasilkan diagnosis. Untuk membangun suatu solusi yang jauh lebih cepat dalam menghasilkan proses keputusan dan memiliki tingkat keakuratan yang tinggi, maka penelitian ini mencoba membangun aplikasi ML dengan menggunakan algoritma SVM. ML dipilih sebagai solusi karena aplikasi yang dibangun dapat memiliki sifat *learning* tanpa harus diprogram berulang-ulang. Dengan *dataset* yang tepat dan lengkap, ML juga akan menghasilkan prediksi yang akurat. Algoritma SVM dipilih karena memiliki nilai akurasi yang jauh lebih baik dibandingkan dengan algoritma lainnya. Penerapan Algoritma SVM dibagi kedalam dua proses, yaitu (1) proses *training* dan (2) proses *testing*. Proses *training* dilakukan untuk menemukan model, dan proses *testing* digunakan untuk membuktikan keakuratan model yang dihasilkan. Pada proses *training* digunakan lima buah data sample untuk menemukan model. Tabel 5, menggambarkan *sample* dari data *training* yang akan digunakan.

Tabel 5. Sample data training

H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11
S1	5	1	1	1	2	1	3	1	1	-1
S2	5	4	4	5	7	10	3	2	1	-1
S3	3	1	1	1	2	2	3	1	1	-1
S4	8	10	10	8	7	10	9	7	1	1
S5	8	7	5	10	7	9	5	5	4	1

Sumber: <http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>

Pada *sample* data *training* nilai H1 diubah menjadi S1 dan seterusnya untuk mempersingkat, dan nilai pada H11 yang sebelumnya bernilai 2 (untuk *class* jinak) dan nilai 4 (untuk *class* ganas) diubah menjadi -1 (untuk *class* jinak) dan 1 (untuk *class* ganas) untuk memudahkan proses perhitungan. Proses *training* memiliki delapan langkah penyelesaian. Berikut ini hasil dari langkah-langkah pada proses *training* yang dilakukan.

(1) Inisiasi nilai

$$\alpha = 0$$

$$C = 1$$

$$\varepsilon = 0,001$$

$$\gamma = 0,5$$

$$\lambda = 0,5$$

(2) Hitung *Dot Product* data training

Tabel 6, menggambarkan hasil perhitungan yang diilustrasikan dalam bentuk matriks.

Tabel 6. Hasil Dot Product data training

	A1	A2	A3	A4	A5
A1	44	74	35	127	109
A2	74	245	74	351	306
A3	35	74	31	121	102
A4	127	351	121	608	487
A5	109	306	102	487	434

(3) Hitung matriks

Perhitungan dimulai DA_1A_1 dari sampai DA_5A_5 dan sehingga menghasilkan nilai matriks yang diilustrasikan pada Tabel 7.

Tabel 7. Hasil matriks

44,25	74,25	35,25	-127,25	-109,25
74,25	245,25	74,25	-351,25	-306,25
35,25	74,25	31,25	-121,25	-102,25
-127,25	-351,25	-121,25	608,25	487,25
-109,25	-306,25	-102,25	487,25	434,25

(4) Cari nilai E (Error)

$$E_1 = -41,375$$

$$E_2 = -131,875$$

$$E_3 = -41,375$$

$$E_4 = 247,875$$

$$E_5 = 201,875$$

(5) Hitung $\Delta\alpha$

$$\Delta\alpha_1 = 0,5$$

$$\Delta\alpha_2 = 0,5$$

$$\Delta\alpha_3 = 0,5$$

$$\Delta\alpha_4 = -0,5$$

$$\Delta\alpha_5 = -0,5$$

(6) Jika $\Delta\alpha < \varepsilon$ maka iterasi selanjutnya

Karena nilai $\Delta\alpha$ kurang dari 0,001 maka itersi dilanjutkan, sehingga dihasilkan nilai sebagai berikut.

$$\Delta\alpha_1 = 1$$

$$\Delta\alpha_2 = 1$$

$$\Delta\alpha_3 = 1$$

$$\Delta\alpha_4 = 0$$

$$\Delta\alpha_5 = 0$$

(7) Mencari nilai w^+ dan w^-

$$w^+ = -599,75$$

$$w^- = 393,75$$

(8) Mencari nilai bias (b)

$$b = 103$$

2. Proses Testing

Pada proses *testing* dibutuhkan sebuah data *sample* untuk menguji model yang telah dihasilkan pada tahapan *training*. Tabel 8, menggambarkan *sample* data untuk proses *testing*.

Tabel 8. Sample data tetsing

H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
Sx	6	8	8	1	3	4	3	7	1

Tahapan *testing* memiliki dua langkah penyelesaian. Berikut ini hasil dari tahapan *testing* yang dilakukan.

(1) Hitung *Dot Product* data *sample* dengan semua data *training*

Pada langkah ini dilakukan perhitungan untuk mendapatkan hasil *Dot Product* dari data *testing*. Tabel 9, menggambarkan hasil *Dot Product* dari data *testing*.

Tabel 9. Hasil Dot Product data testing

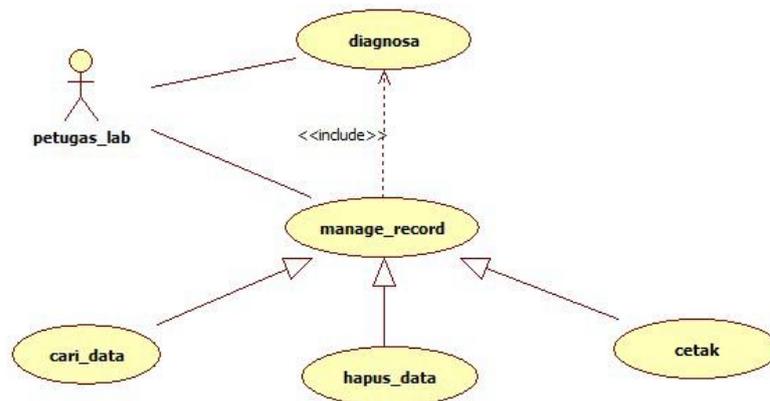
A1	A2	A3	A4	A5
74	184	66	354	265

(2) Hitung keputusan

Hasil perhitungan keputusan adalah $F(\mathbf{x}) = -221$, karena hasilnya bernilai negatif maka data sample tersebut masuk dalam kelas **-1** yaitu jinak.

3. Perancangan dan Implementasi

Proses perancangan aplikasi digunakan diagram use case untuk menggambarkan interaksi aktor dengan aplikasi. Gambar 1, menggambarkan use case diagram pada aplikasi ML diagnosis penyakit kanker payudara menggunakan algoritma SVM.



Gambar 1: Use Case Diagram

Hasil implementasi aplikasi ML diagnosis penyakit kanker payudara menggunakan algoritma SVM, ditampilkan pada Gambar 2, 3 dan 4, berikut ini.



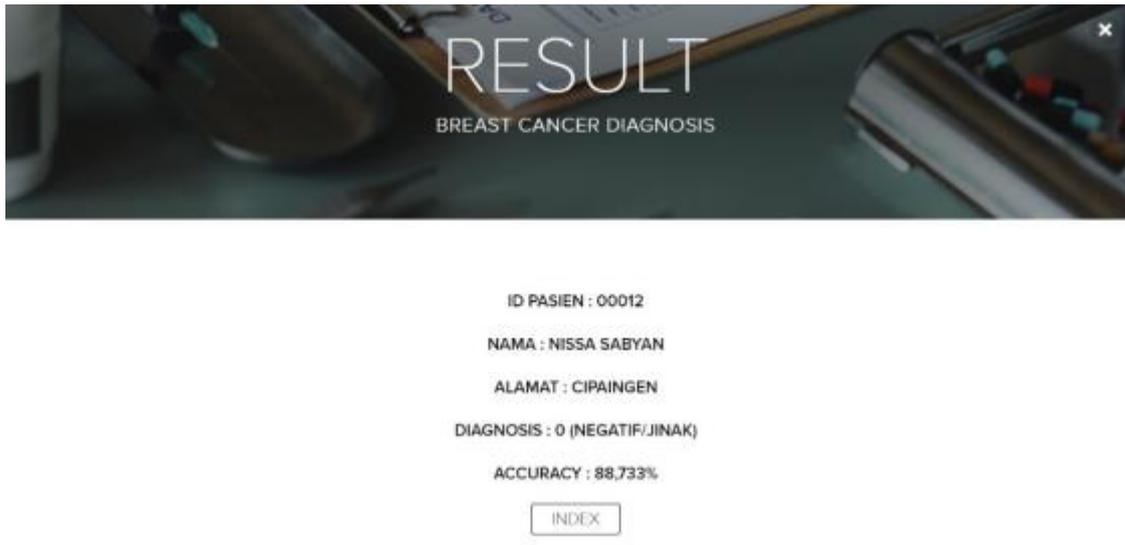
DIAGNOSIS

MANAGE RECORD

Gambar 2: Tampilam halaman awal aplikasi

A screenshot of the 'Breast Cancer Diagnosis' application interface. The form is titled 'Breast Cancer Diagnosis' and is divided into two main sections. The first section, 'Profil Pasien', contains two input fields for 'Nama' and 'Alamat'. The second section, 'Hasil Lab', contains nine input fields arranged in a 3x3 grid. The labels for these fields are: 'Clump thickness', 'Marg adhesion', 'Bland chrom', 'Unif cell size', 'Single epith cell size', 'Norm nucleoli', 'Unif cell shape', 'Bare nuclei', and 'Mitoses'. A blue 'Diagnosa' button is located at the bottom left of the form.

Gambar 3: Tampilam halaman input diagnosis



Gambar 4: Tampilan halaman hasil diagnosis

4. SIMPULAN

Hasil penelitian menunjukkan bahwa aplikasi ML untuk mendiagnosis penyakit kanker payudara dapat menghasilkan prediksi keputusan berdasarkan dua kemungkinan, yaitu sel hidup dalam kondisi ganas atau jinak. Aplikasi ini dapat digunakan oleh petugas lab untuk melakukan diagnosis berdasarkan hasil FNA yang diperoleh dari pemeriksaan biopsi. Penggunaan algoritma SVM pada aplikasi ML dapat meningkatkan keakuratan hasil diagnosa menjadi lebih tinggi, karena semakin banyak data yang bersifat lengkap dan tepat maka akan menambah tingkat keakuratan dari hasil diagnosisnya. Berdasarkan penelitian yang telah dilakukan, maka beberapa hal penting yang perlu diperhatikan dalam membangun aplikasi ML, adalah sebagai berikut, (1) dataset dan (2) algoritma untuk mencari pola data dan menentukan prediksi. Dataset yang lengkap dan tepat sangat berpengaruh terhadap hasil prediksi. Pemilihan algoritma yang tepat pula dapat menunjang keakuratan hasil prediksi. Pada penelitian selanjutnya dapat dikembangkan dengan menggunakan model algoritma dan dataset lainnya sebagai bahan uji coba untuk menentukan model terbaik dalam membangun aplikasi ML.

5. DAFTAR PUSTAKA

- Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Computer Science*, 83(Fams), 1064–1069.
<https://doi.org/10.1016/j.procs.2016.04.224>
- Edition, S. (2017). *AJCC Cancer Staging Manual. AJCC Cancer Staging Manual*.
<https://doi.org/10.1007/978-3-319-40618-3>
- Edriss, E., Ali, E., & Feng, W. Z. (2016). Breast Cancer Classification using Support Vector Machine and Neural Network. *International Journal of Science and Research (IJSR)*, 5(3), 1–6. <https://doi.org/10.21275/v5i3.nov161719>
- Janardhanan, P., Heena, L., & Sabika, F. (2015). Effectiveness of support vector machines in medical data mining. *Journal of Communications Software and Systems*, 11(1), 25–30. <https://doi.org/10.24138/jcomss.v11i1.114>
- Primartha, R (2018). Belajar Machine Learning. Informatika. Bandung.
- Son, Y. J., Kim, H. G., Kim, E. H., Choi, S., & Lee, S. K. (2010). Application of support vector machine for prediction of medication adherence in heart failure patients. *Healthcare Informatics Research*, 16(4), 253–259.
<https://doi.org/10.4258/hir.2010.16.4.253>
- W. N. Street, O. L. Mangasarian, and W.H. Wolberg. Breast Cancer Wisconsin (Prognostic) Dataset. UCI.
<http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/>
- Zhao, C. Y., Zhang, H. X., Zhang, X. Y., Liu, M. C., Hu, Z. D., & Fan, B. T. (2006). Application of support vector machine (SVM) for prediction toxic activity of different data sets. *Toxicology*, 217(2–3), 105–119.
<https://doi.org/10.1016/j.tox.2005.08.019>