

**DIAGNOSA PENDERITA PENYAKIT KANKER PAYUDARA
MENGUNAKAN METODE NAÏVE BAYES**



SKRIPSI

Diajukan untuk memenuhi salah satu syarat kelulusan Program Sarjana

TAUFIK FRISSETYO

11180817

**STMIK
NUSA MANDIRI**

Program Studi Sistem Informasi

STMIK Nusa Mandiri

Jakarta

2019

ABSTRAK

Taufik Frissetyo (11180817), Diagnosa Penderita Penyakit Kanker Payudara Menggunakan Metode *Naïve Bayes*

Perkembangan teknologi informasi yang begitu cepat mempengaruhi lini kehidupan manusia. Hal ini dibuktikan dengan diciptakannya sistem berbasis komputer yang dapat mengatasi persoalan disegala lini kehidupan. Salah satu bidang yang saat ini dipengaruhi oleh teknologi informasi adalah bidang kesehatan dan kedokteran. Penyakit kanker semakin hari semakin banyak diderita oleh sebagian besar orang bahkan tanpa mereka sadari. Hal ini diakibatkan karena pola hidup yang kurang sehat dan ada sebagian orang yang sudah tidak terlalu peduli akan kesehatannya padahal kesehatan merupakan hal yang amat sangat penting bagi kehidupan manusia. Salah satu Kanker yang paling banyak diderita oleh masyarakat khususnya kaum perempuan dan sangat mematikan adalah Kanker Payudara. Maka perlunya deteksi dini kanker sangatlah penting demi keselamatan penderita. Metode yang digunakan penulis dalam penelitian ini adalah metode *Naïve Bayes* karena metode ini sangat sederhana tetapi memiliki tingkat akurasi yang baik. Hasil perhitungan yang penulis lakukan baik secara manual dan menggunakan *Rapidminer* diketahui akurasi untuk data *training* sebesar 81.00% dan *AUC* sebesar 0,922 dengan kategori *Excellent Classification* dan hasil akurasi data *testing* sebesar 62.50% dan *AUC* sebesar 0,719 dengan kategori *Fair Classification*.

Kata Kunci : Diagnosa, Kanker, Payudara, *Naïve Bayes*

STMIK
NUSA MANDIRI

ABSTRACT

Taufik Frissetyo (11180817), Diagnosis of Patients with Breast Cancer Using the Naïve Bayes Method

The development of information technology is so fast influencing the lines of human life. This is evidenced by the creation of computer-based systems that can overcome problems in all lines of life. One of the fields currently affected by information technology is health and medicine. Cancer is increasingly suffering by most people even without them knowing it. This is caused by unhealthy lifestyles and there are some people who are not too concerned about their health even though health is very very important for human life. One of the most common cancers suffered by the community, especially women and is very deadly is Breast Cancer. So the need for early detection of cancer is very important for patient safety. The method used by the authors in this study is the Naïve Bayes method because this method is very simple but has a good level of accuracy. The results of calculations that the authors do both manually and using Rapidminer known accuracy for training data of 81.00% and AUC of 0.922 with the category of Excellent Classification and the results of data testing accuracy of 62.50% and AUC of 0.719 with the Fair Classification category.

Keywords : Diagnosis, Cancer, Breast, Naïve Bayes



DAFTAR ISI

LEMBAR JUDUL SKRIPSI.....	i
LEMBAR PERSEMBAHAN.....	ii
LEMBAR PERNYATAAN KEASLIAN SKRIPSI.....	iii
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH .	iv
LEMBAR PERSETUJUAN DAN PENGESAHAN SKRIPSI.....	v
LEMBAR PANDUAN PENGGUNAAN HAK CIPTA.....	vi
 Kata Pengantar	vii
Abstrak.....	ix
Daftar Isi	xi
Daftar Gambar.....	xiii
Daftar Tabel.....	xiv
Daftar Lampiran.....	xv
 BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Identifikasi Permasalahan.....	3
1.3. Maksud dan Tujuan.....	3
1.4. Metode Penelitian.....	4
A. Observasi	4
B. Studi Pustaka	4
1.5. Ruang Lingkup.....	4
1.6. Hipotesa	4
 BAB II LANDASAN TEORI	5
2.1. Tinjauan Pustaka	5
2.1.1. Data Mining.....	5
2.1.2. Klasifikasi	8
2.1.3. Kanker.....	9
2.1.4. Payudara	10
2.1.5. Naïve Bayes.....	12
2.2. Penelitian Terkait.....	14
2.3. Tinjauan Objek Penelitian.....	15
 BAB III METODOLOGI PENELITIAN.....	19
3.1. Tahapan Penelitian.....	19
3.2. Instrumen Penelitian	20
3.3. Metode Pengumpulan Data, Populasi dan Sampel Penelitian.....	21
3.4. Metode Analisis Data	21
 BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	23
4.1. Pengumpulan Dataset.....	23
4.2. Pengolahan Data Awal.....	23
4.3. Transformasi Data.....	24
4.4. Menentukan Metode Yang Digunakan.....	26

4.5. Eksperimen dan Pengujian Metode.....	27
4.6. Evaluasi.....	37
BAB V PENUTUP	41
5.1. Kesimpulan	41
5.2. Saran.....	41

DAFTAR PUSTAKA
DAFTAR RIWAYAT HIDUP
LEMBAR KONSULTASI BIMBINGAN
LAMPIRAN



DAFTAR PUSTAKA

- Agustina, D. M., & Wijanarto. (2016). Analisis Perbandingan Algoritma ID3 Dan C4 . 5 Untuk Klasifikasi Penerima Hibah Pemasangan Air Minum Pada PDAM Kabupaten Kendal. *Journal of Applied Intelligent System*, 1(3), 234–244.
- Argario, H. B., Hidayat, N., & Dewi, R. K. (2018). Implementasi Metode Naive Bayes Untuk Diagnosis Penyakit Kambing (Studi Kasus : UPTD . Pembibitan Ternak dan Hijauan Makanan Ternak Kec. Singosari Malang). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 2(8), 2719–2723.
- Ervinaeni, Y., Hidayat, A. S., & Riana, E. (2019). Sistem Pakar Diagnosa Gangguan Hiperaktif Pada Anak Dengan Metode Naive Bayes Berbasis Web. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 3(2), 90–104. <https://doi.org/10.30865/mib.v3i2.1158>
- Farahdiba, B. A., & Nugroho, Y. S. (2016). Klasifikasi Kanker Payudara Menggunakan Algoritma Gain Ratio. *Jurnal Teknik Elektro*, 8(2), 43–46.
- Haryono, S. J., Anwar, S. L., & Salim, A. (2018). *Dasar-Dasar Biologi Molekul Kanker bagi Praktisi Klinis*. Yogyakarta: ugmppress.
- Kurniasari, F. N., Harti, L. B., Ariestiningsih, A. D., Wardhani, S. O., & Nugroho, S. (2017). *Buku Ajar Gizi dan Kanker*. Malang: UB Press.
- Kurniawan, M. F., & Ivandari. (2017). KOMPARASI ALGORITMA DATA MINING UNTUK KLASIFIKASI PENYAKIT KANKER PAYUDARA. *IC-Tech*, 12(1), 1–8.
- Ma'arif, F., & Arifin, T. (2017). Optimasi Fitur Menggunakan Backward Elimination Dan Algoritma SVM Untuk Klasifikasi Kanker Payudara. *Jurnal Informatika (JI) UBSI*, 4(1), 46–53. <https://doi.org/10.31311/JI.V4I1.1548>
- Moriesta, E., Selviani, & Ibrahim, A. (2017). Analisis Penyaringan Email Spam Menggunakan Metode Naive Bayes. *Prosiding Annual Research Seminar 2017*, 3(1), 45–48.
- Muflikhah, L., Ratnawati, D. E., & Putri, R. R. M. (2018). *Data Mining*. Malang: UB Press.
- Naviri, T. (2016). *Buku Pintar Kesehatan dan Kecantikan Payudara*. Jakarta: Elex Media Komputindo.
- Nofriansyah, D., & Nurcahyo, G. W. (2015). *Algoritma Data Mining dan Pengujian*. Sleman: deepublish.
- Praningki, T., & Budi, I. (2018). Sistem Prediksi Penyakit Kanker Serviks Menggunakan CART, Naive Bayes, dan k-NN. *Creative Information Technology Journal*, 4(2), 83–93. <https://doi.org/10.24076/citec.2017v4i2.100>

Rifai, M. H., & Wijayanti, A. (2017). PEMANFAATAN MEDIA PEMBELAJARAN GEOGRAFI SMA DI KABUPATEN KARANGANYAR. *Jurnal Edudikara*, 2(3), 210–216.

Sawari, R. C. (2018). Ciri-ciri Kanker Payudara yang Masih Bisa Disembuhkan. Retrieved May 5, 2019, from Hallo Sehat website: <https://hellosehat.com/hidup-sehat/tips-sehat/ciri-ciri-kanker-payudara-stadium-awal/>

Septiani, W. D. (2017). KOMPARASI METODE KLASIFIKASI DATA MINING ALGORITMA C4.5 DAN NAIVE BAYES UNTUK PREDIKSI PENYAKIT HEPATITIS. *Jurnal Pilar Nusa Mandiri*, 13(1), 76–84.

Syahputra, R., & Safitri, W. (2018). Analisa Tingkat Prestasi Atlet Karate Sumatera Barat Menggunakan Algoritma Data Mining. *Indonesian Journal of Computer Science*, 7(2), 200–210.

Via, Y. V., Nugroho, B., & Syafrizal, A. (2015). Sistem Pendukung Keputusan Klasifikasi Tingkat Keganasan Kanker Payudara Dengan Metode Naive Bayes Classifier. *SCAN-Jurnal Teknologi Informasi Dan Komunikasi*, 10(2), 63–68.

Wahyuni, E. S. (2016). PENERAPAN METODE SELEKSI FITUR UNTUK MENINGKATKAN HASIL DIAGNOSIS KANKER PAYUDARA. *Jurnal SIMETRIS*, 7(1), 283–294.

Widayu, H., Darma, S., Silalahi, N., & Mesran. (2017). Data Mining Untuk Memprediksi Jenis Transaksi Nasabah Pada Koperasi Simpan Pinjam Dengan Algoritma C4.5. *MEDIA INFORMATIKA BUDIDARMA*, 1(1), 32–37.

Yanto, R., & Khoiriah, R. (2015). Implementasi Data Mining dengan Metode Algoritma Apriori dalam Menentukan Pola Pembelian Obat. *Creative Information Technology Journal*, 2(2), 102–113. <https://doi.org/10.24076/citec.2015v2i2.41>

Zulfikar, W. B., & Lukman, N. (2017). Perbandingan Naive Bayes Classifier Dengan Nearest Neighbor Untuk Identifikasi Penyakit Mata. *Jurnal Online Informatika*, 1(2), 82–86. <https://doi.org/10.15575/join.v1i2.33>



**STMIK
NUSA MANDIRI**