

Implementasi Algoritma Random Forest Untuk Mendiagnosis Kejadian Berulang (Kekambuhan) Pada Kanker Payudara Berbasis Web

Runi Hari Bagus Saputra^{1*}, Roy Mubarak²

¹Teknik, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: ¹runihari Bagus171@gmail.com , ²dosen02685@unpam.ac.id

(*: Corresponden Author)

Abstrak– Kanker merupakan penyebab kematian utama di negara maju dan penyebab kematian kedua di negara berkembang. Kementerian Kesehatan Republik Indonesia (2019) menyatakan bahwa kejadian kanker di Indonesia (penduduk 136,2/100.000) berada di urutan ke-8 di Asia Tenggara dan ke-23 di Asia, menurut data stasiun kanker global pada tahun 2018 meningkat. Insiden tertinggi pada wanita, dengan kanker payudara sebesar 42,1 per 100.000 penduduk dan angka kematian rata-rata 17 per 100.000 penduduk, diikuti oleh kanker serviks sebesar 23,4 per 100.000 penduduk, dengan angka kematian rata-rata 13,9 per 100.000 penduduk. Dengan anggapan bahwa kanker payudara telah hilang dari tubuh penderita, tidak banyak orang yang mengetahui bahwa kanker payudara dapat kambuh lagi. Setelah proses pengobatan, biasanya hanya beberapa sel kanker yang bertahan. Tetapi pada dosis kecil, sel-sel ini dapat menyebabkan masalah yang lebih serius. Penyakit ini bisa kambuh di dada, dada, atau bagian tubuh lainnya, seperti tulang dan hati. Untuk memprediksi kekambuhan kanker payudara, diperlukan suatu sistem berbasis teknologi yang dapat memprediksi kekambuhan kanker payudara. Algoritma yang digunakan adalah Random Forest. Ini adalah kombinasi dari semua pohon yang sesuai dalam satu mode. Sistem ini berbasis bahasa pemrograman PHP dan database MySQL sebagai media penyimpanannya. Bahasa pemrograman PHP adalah bahasa pemrograman yang umum digunakan untuk membuat dan mengembangkan situs web dan dapat digunakan dalam kombinasi dengan HTML. MySQL, di sisi lain, adalah "manajer database" dari sistem manajemen database. Ini menggunakan perintah dasar SQL (Structured Query Language).

Kata Kunci : *Diagnosis Kejadian Berulang, Kanker Payudara, Random Forest*

Abstract– Cancer is the leading cause of death in developed countries and the second leading cause of death in developing countries. The Ministry of Health of the Republic of Indonesia (2019) stated that the incidence of cancer in Indonesia (population 136.2/100,000) was 8th in Southeast Asia and 23rd in Asia, according to data from the global cancer station in 2018. The highest incidence is in women, with breast cancer at 42.1 per 100,000 population and an average mortality rate of 17 per 100,000 population, followed by cervical cancer at 23.4 per 100,000 population, with an average mortality rate of 13.9 per 100,000 population. . With the assumption that breast cancer has disappeared from the patient's body, not many people know that breast cancer can recur. After the treatment process, usually only a few cancer cells survive. But at small doses, these cells can cause more serious problems. This disease can recur in the chest, chest, or other body parts, such as the bones and liver. To predict the recurrence of breast cancer, we need a technology-based system that can predict the recurrence of breast cancer. The algorithm used is Random Forest. It is a combination of all matching trees in one mode. This system is based on the PHP programming language and MySQL database as the storage medium. The PHP programming language is a programming language that is commonly used to create and develop websites and can be used in combination with HTML. MySQL, on the other hand, is the "database manager" of the database management system. It uses basic SQL (Structured Query Language) commands.

Keywords : *Diagnosis of Recurrent Occurrence, Breast Cancer, Random Forest*

1. PENDAHULUAN

Kanker merupakan penyakit dengan penyebab utama kematian pada negara dengan ekonomi maju dan penyebab kematian kedua pada negara dengan ekonomi berkembang. Kemenkes RI (2019) mengemukakan, *Data Global Cancer Observatory* Tahun 2018 menunjukkan angka kejadian penyakit kanker di Indonesia (136,2/100.000 penduduk) berada pada urutan 8 di Asia Tenggara, sedangkan di Asia urutan ke 23. Angka kejadian untuk perempuan yang tertinggi adalah kanker payudara yaitu sebesar 42,1 per 100.000 penduduk dengan rata-rata kematian 17 per 100.000 penduduk yang diikuti kanker leher rahim sebesar 23,4 per 100.000 penduduk dengan rata-rata kematian 13,9 per 100.000 penduduk.

Dengan anggapan bahwa kanker payudara telah hilang dari tubuh penderita, tidak banyak orang yang mengetahui bahwa kanker payudara dapat kambuh lagi. Setelah proses pengobatan, biasanya hanya beberapa sel kanker yang bertahan. Tetapi pada dosis kecil, sel-sel ini dapat menyebabkan masalah yang lebih serius. Penyakit ini bisa kambuh di dada, dada, atau bagian tubuh lainnya, seperti tulang dan hati. (GueSehat, 2017).

Untuk memprediksi kekambuhan kanker payudara, diperlukan suatu sistem berbasis teknologi yang dapat memprediksi kekambuhan kanker payudara. Algoritma yang digunakan adalah Random Forest. Ini adalah kombinasi dari semua pohon yang cocok dan digabungkan menjadi satu mode (aditya.yanuar.r., 2018). Sistem ini berbasis bahasa pemrograman PHP dan database MySQL sebagai media penyimpanannya. Bahasa pemrograman PHP adalah bahasa pemrograman yang umum digunakan untuk membuat dan mengembangkan situs web dan dapat digunakan dalam kombinasi dengan HTML. MySQL, di sisi lain, adalah "manajer database" dari sistem manajemen database. Ini menggunakan perintah dasar SQL (Structured Query Language).

2. METODOLOGI PENELITIAN

Dalam penelitian ini menggunakan dua metode yaitu metode pengumpulan data dan metode pengembangan sistem

a. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah Metode Studi Pustaka, yaitu teknik mencari data yang berkaitan dengan penelitian yang bersumber dari buku-buku dan dokumen-dokumen berupa jurnal, skripsi serta data-data terkait yang diperlukan untuk menunjang pembangunan sistem.

b. Metode Pengembangan Sistem

Pengembangan sistem merupakan penyusunan suatu sistem yang baru untuk menggantikan sistem yang lama secara keseluruhan atau memperbaiki sistem yang telah ada. Maka pengembangan yang sesuai untuk sistem yang dibuat yaitu menggunakan metode *Rapid Application Development* (RAD).

3. ANALISA DAN PERANCANGAN

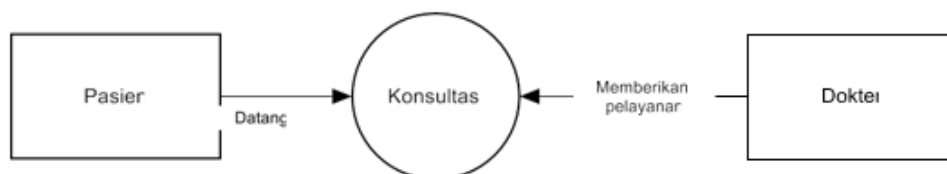
3.1 Analisa Sistem

3.1.1 Analisa Sistem Saat Ini

Analisis sistem saat ini merupakan gambaran tentang sistem yang saat ini sedang berjalan dalam mendiagnosis kekambuhan pada kanker payudara.

1) Prosedur Sistem Berjalan

Prosedur sistem berjalan adalah suatu proses atau langkah-langkah yang menunjukkan perjalanan data atau dokumen pada suatu sistem atau proses keluar masuknya berkas yang terjadi pada suatu sistem yang sedang berjalan atau berlangsung. Tujuan dari penulisan ini adalah untuk mengetahui gambaran secara jelas mengenai sistem yang sedang berjalan dalam mendiagnosis kekambuhan pada kanker payudara.



Gambar 1. Diagram Konteks Sistem Berjalan

3.1.2 Analisa Data

Analisa data merupakan suatu langkah yang paling menentukan dari suatu penelitian, karena analisa data berfungsi untuk menyimpulkan hasil penelitian. Analisis data dapat dilakukan melalui tahap berikut ini:

- a. Tahap Penelitian
 1. Perencanaan

Pada tahap ini kegiatan yang dilakukan adalah sebagai berikut:

 - a) Peneliti merancang kelas yang akan dijadikan sampel.
 - b) Peneliti membuat instrumen-instrumen penelitian yang akan digunakan untuk penelitian
 2. Pelaksanaan

Pada tahap ini kegiatan yang dilakukan adalah sebagai berikut:

 - a) Peneliti melaksanakan pembelajaran pada sampel penelitian.
 - b) Peneliti menguji coba, menganalisis dan menetapkan instrumen penelitian
 3. Evaluasi

Pada tahap ini, peneliti menganalisis dan mengolah data yang telah dikumpulkan dengan metode yang telah ditentukan
 4. Penyusunan Laporan

Pada tahap ini, kegiatan yang dilakukan adalah menyusun dan melaporkan hasil-hasil penelitian
- b. Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini berupa data kanker payudara yang diambil dari *Mechine Learning Repository* (sumber: <https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer/>). Data latih yang digunakan terlihat pada Tabel 1

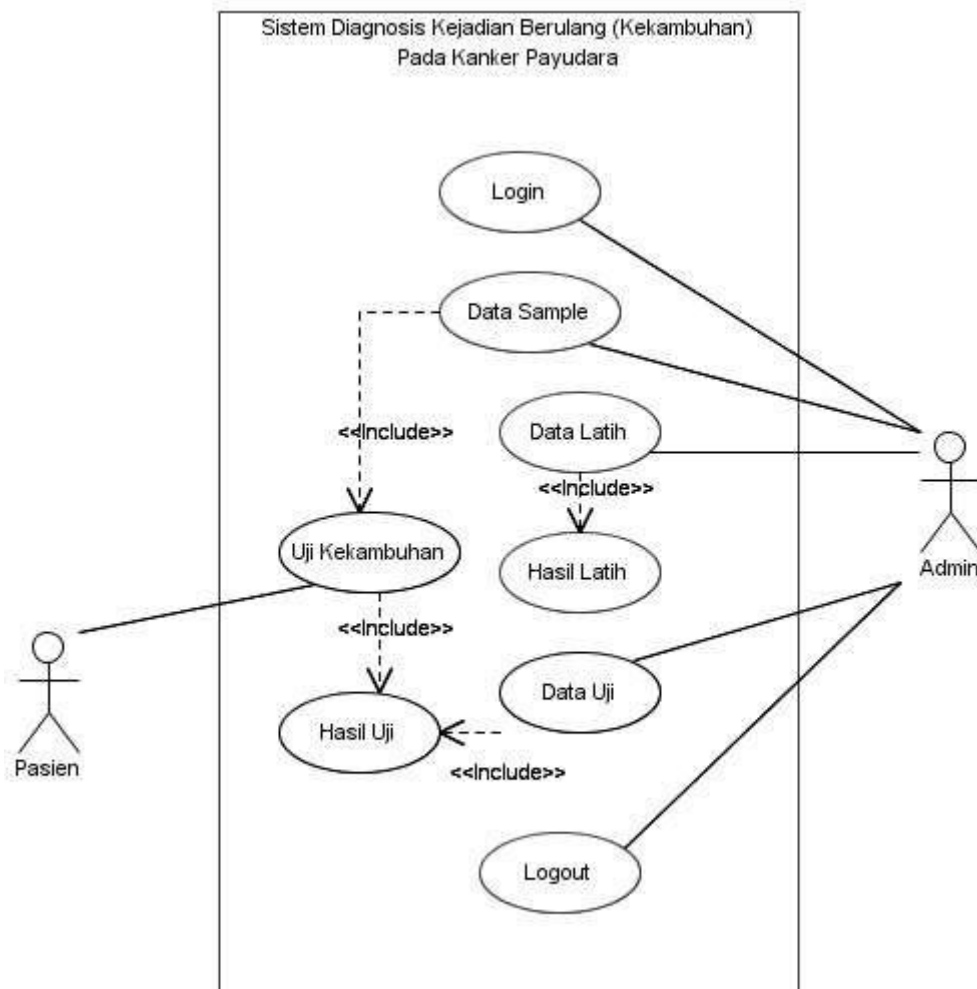
Tabel 1 Data Latih

N o.	Class	Age	Menopause	Tumor-size	Inv-nodes	Code-caps	De-gmalig	Breast	Breast-quad	Irradiat
1	recurrence-events	50-59	premeno	30-34	0-2	no	3	right	left_up	yes
2	no-recurrence-events	50-59	premeno	30-34	3-5	yes	2	left	left_low	yes
3	no-recurrence-events	50-59	ge40	35-39	0-2	no	3	left	left_low	no
4	no-recurrence-events	50-59	ge40	35-39	0-2	no	3	left	left_up	no
5	no-recurrence-events	50-59	ge40	35-39	0-2	no	2	left	left_up	no
6	recurrence-events	50-59	ge40	35-39	0-2	no	2	left	left_low	no

7	no-recurrence-events	50 - 59	ge40	35-39	15-17	no	3	left	left_lo w	no
8	recurrence-events	60 - 69	ge40	35-39	6-8	yes	3	left	left_lo w	no
9	no-recurrence-events	60 - 69	ge40	45-49	6-8	yes	3	left	centra l	no
10	no-recurrence-events	40 - 49	premeno	45-49	0-2	no	2	left	left_lo w	yes
11	no-recurrence-events	50 - 59	ge40	50-54	0-2	no	1	right	right_ up	no
12	no-recurrence-	60 - 69	ge40	50-54	0-2	no	2	left	left_lo w	no
13	no-recurrence-events	60 - 69	ge40	50-54	0-2	no	2	right	left_u p	yes
14	recurrence-events	60 - 69	ge40	50-54	0-2	no	3	right	left_u p	no
15	no-recurrence-events	40 - 49	premeno	50-54	0-2	no	2	left	left_lo w	no
16	recurrence-events	40 - 49	premeno	50-54	0-2	no	2	right	left_lo w	yes
17	no-recurrence-events	50 - 59	premeno	50-54	0-2	yes	2	right	left_u p	yes
18	recurrence-events	50 - 59	premeno	50-54	9-11	yes	2	right	left_u p	no
19	no-recurrence-events	50 - 59	ge40	5-9	0-2	no	2	right	right_ up	no
20	no-recurrence-events	60 - 69	ge40	5-9	0-2	no	1	left	central	no
21	no-recurrence-events	30 - 39	premeno	5-9	0-2	no	2	left	right_l o w	no
22	no-recurrence-events	40 - 49	premeno	5-9	0-2	no	1	left	left_lo w	yes
23	no-recurrence-events	50 - 59	ge40	0-4	0-2	no	2	left	central	no
24	no-recurrence-events	50 - 59	ge40	0-4	0-2	no	1	right	central	no
25	no-recurrence-	50 -	ge40	0-4	0-2	no	1	left	left_lo w	no

	events	59								
26	no-recurrence-events	70 - 79	ge40	0-4	0-2	no	1	left	right_low	no
27	no-recurrence-events	30 - 39	premeno	0-4	0-2	no	2	right	central	no
28	recurrence-events	30 - 39	premeno	0-4	0-2	no	2	right	central	no
29	no-recurrence-events	40 - 49	premeno	0-4	0-2	no	2	right	right_low	no
30	no-recurrence-events	40 - 49	premeno	0-4	0-2	no	3	left	central	no

3.1.3 Use Case Diagram



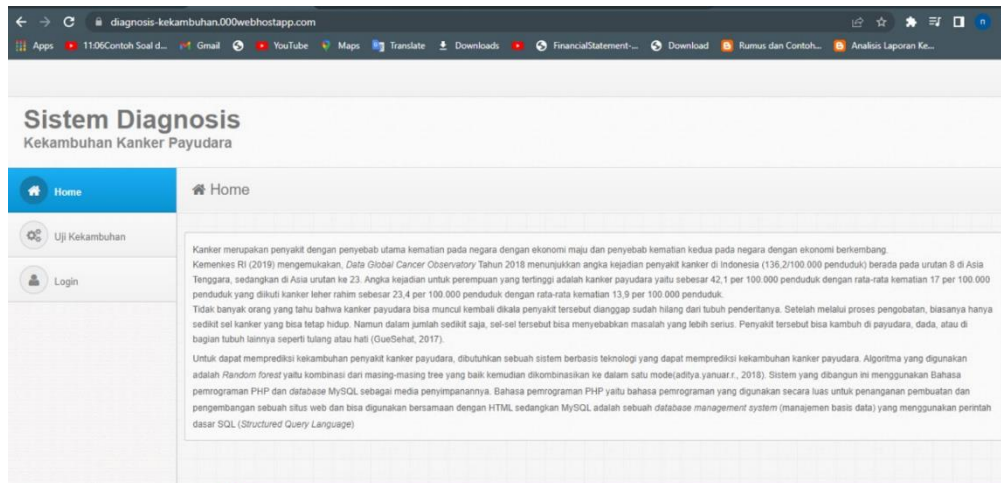
Gambar 2. Usecase Diagram

Gambar diatas merupakan gambaran dari usecase diagram dari awal login hingga logout.

4. IMPLEMENTASI

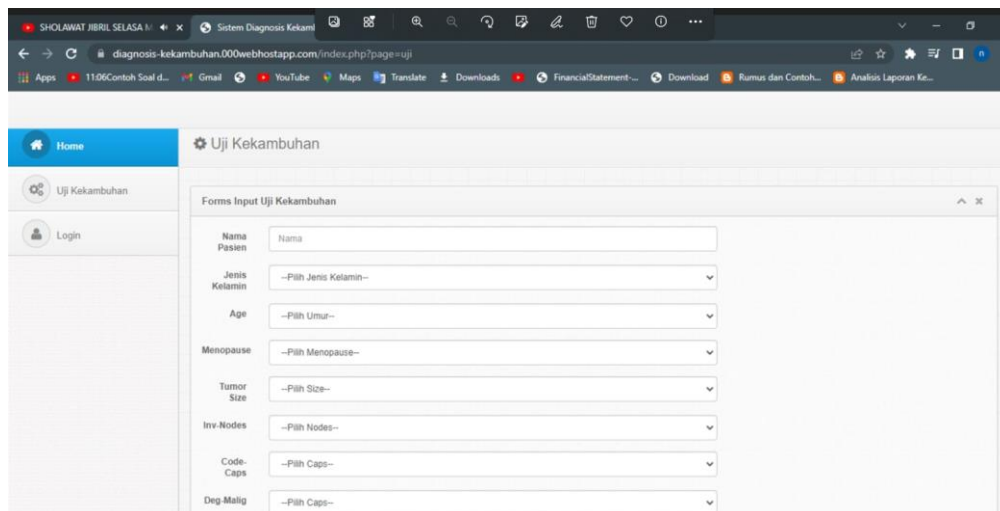
Sistem Diagnosis Kekambuhan Kanker Payudara telah implementasi dan di publish, situs dapat dibuka melalui URL <https://diagnosis-kekambuhan.000webhostapp.com/>. Berikut ini adalah hasil *screen shoot* tampilan pada situs:

a. Halaman Login

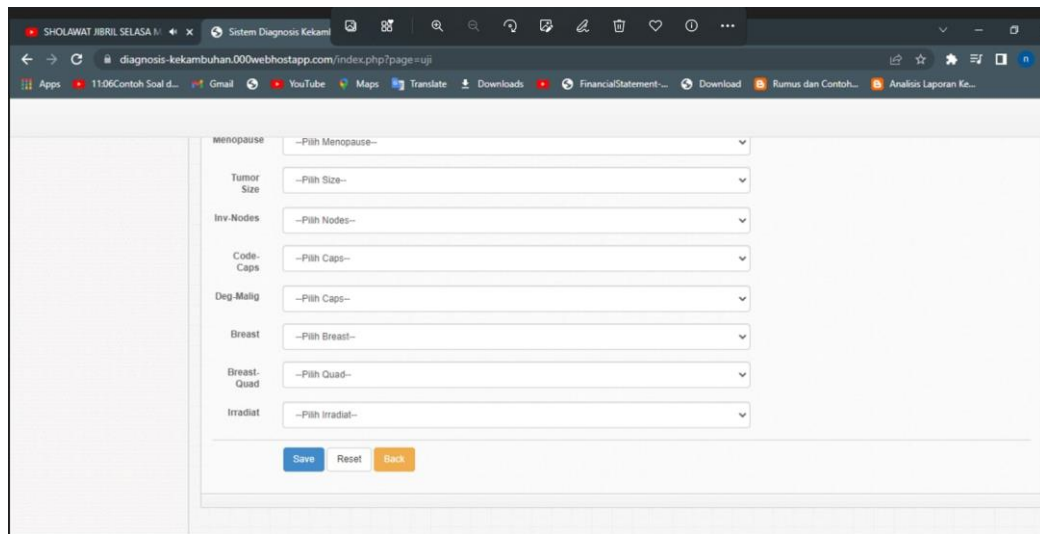


Gambar 3. *Screen shoot* Tampilan Halaman Home

b. Halaman Form Input Uji Kekambuhan



Gambar 4. *Screen shoot* Tampilan Halaman Form Input Uji Kekambuhan



Gambar 5. Screen shoot Tampilan Halaman Form Input Uji Kekambuhan

5. KESIMPULAN

a. Kesimpulan

Berikut kesimpulan yang dapat penulis uraikan mengenai Sistem Diagnosis Kekambuhan Kanker Payudara berbasis web bahwa kekambuhan pada kanker payudara dapat di diagnosis dengan menggunakan sistem pakar. Implementasi Algoritma Random Forest untuk mendiagnosis kekambuhan kanker payudara dengan cara membagi setiap simpul menjadi variable berdasarkan variable terbaik diantara variable yang dipilih secara acak disetiap simpul.

b. Saran

Saran yang dapat peneliti sampaikan untuk penelitian selanjutnya yaitu:Melakukan komparasi terhadap algoritma atau metode data mining lainnya dalam menganalisa data kanker payudara, untuk mengetahui algoritma mana yang lebih kuat, efisien dan akurat. Sehingga dapat ditentukan algoritma yang tepat yang dapat digunakan dalam menganalisa data kanker payudara. Dapat menggunakan tools lain yakni Rapid Miner, Matlab, dan lain sebagainya untuk mendapatkan analisis yang lebih akurat lagi.

REFERENCES

- Agung Mulyo Widodo, Nizirwan Anwar, Bambang Irawan, Lista Meria, Andika Wisnujati. 2021. "Komparasi Performansi Algoritma Pengklasifikasi KNN, Bagging dan Random Forest Untuk Prediksi Kanker Payudara". Konferensi Nasional Ilmu Komputer (KONIK) 2021. P-ISSN: 2338-2899, E-ISSN: 2807-1271.
- Ahmad Fauzi, Riki Supriyadi dan Nurlaelatul Maulidah. 2020. "Deteksi Penyakit Kanker Payudara dengan Seleksi Fitur berbasis Principal Component Analysis dan Random Forest". Jurnal Infotech Volume 2 No. 1 Juni 2020, E-ISSN: 2715-8160.
- Amalialaisa (Amalia Laisa), Abdullatif (Abdul Latif), Vinashaw (Vina Shaw), & Rinagunawan (Rina Gunawan). (2017, August 17). Apa yang dimaksud dengan Teori Biaya Transaksi atau Transaction Cost Theory?. Sumber: <https://www.dictionio.id/t/apa-yang-dimaksud-dengan-teori-biaya-transaksi-atau-transaction-cost-theory/8996/2> diakses pada 28 September 2021.
- Arizona, N. D. (2017). Aplikasi Pengolahan Data Anggaran Pendapatan dan Belanja Desa (APBDES) Pada Kantor Desa Bakau Kecamatan Jawai Berbasis Web, 01 (02), 105–119.Sumber: <http://openjurnal.unmuhpnk.ac.id/index.php/CN/article/view/745> diakses pada 28 September 2021.
- Balqis Aisyah Farahdiba dan Yusuf Sulisty Nugroho. 2016. "Klasifikasi Kanker Payudara Menggunakan Algoritma Gain Ratio". ISSN 1411 – 0059. Klasifikasi Kanker Payudara Menggunakan Algoritma Gain Ratio.

- Dewaweb Team. 2019. SQL: Pengertian, Fungsi Beserta Perintah Dasarnya. Sumber: <https://www.dewaweb.com/blog/sql-pengertian-fungsi-beserta-perintah-dasarnya/> diakses pada 28 September 2021.
- Dr.Rizal Fadli. Pengertian Kanker Payudara, Gejala Kanker Payudara pada. Sumber: <https://www.halodoc.com/kesehatan/kanker-payudara>. Diakses pada 29 September 2021.
- Dzulhaq, M.Iqbal, Rahmat Tullah, Putra Satia Nugraha 2017. Sistem Informasi Akademik Sekolah Berbasis Kurikulum 2013. Jurnal Sisfotek Global Vol.7, No.1.
- GueSehat. 2017. “Bagaimana Kanker Payudara Bisa Kambuh Lagi?”. Sumber: <https://www.guesehat.com/bagaimana-kanker-payudara-bisa-kambuh-lagi>. Diakses pada 29 September 2021.
- H. Harafani dan H. Aji Al-Kautsar. 2021. “Meningkatkan Kinerja K-NN Untuk Klasifikasi Kanker Payudara Dengan Seleksi Fitur”. Jurnal Pendidikan Teknologi dan Kejuruan Vol. 18, No. 1, Januari 2021, P-ISSN: 0216-3241 E-ISSN: 2541-0652.
- Habibie, R., Isnanto, R. R., & Kridalukmana, R. (2016). Aplikasi Data Pelanggan Berbasis Java Dan Mobile Pada Pt. Pln (Persero) Area Semarang. Jurnal Teknologi Dan Sistem Komputer, 4(1), 50–59 ISSN: 2338-0403.
- Hardian Oktavianto dan Rahman Puji Handri. 2020. “Analisis Klasifikasi Kanker Payudara Menggunakan Algoritma Naive Bayes”. Jurnal Informatika, Vo. 8 No. 2 Februari 2020, ISSN 2301-5632.
- Isty, I., & Afifah, N. (2018). Sistem Informasi Penjualan Busana Pengantin Pada Tutut Manten Yogyakarta, 10(1), 1–6.
- KEMENKES RI. 2019. “Penyakit Kanker di Indonesia Berada Pada Urutan 8 di Asia Tenggara dan Urutan 23 di Asia”. Sumber: <http://p2p.kemkes.go.id/penyakit-kanker-di-indonesia-berada-pada-urutan-8-di-asia-tenggara-dan-urutan-23-di-asia/>. Diakses pada 29 September 2021.
- Kurniawan, Andre. 2021. Analisis Proses Invoicing Account Receivable Dari Transaksi Business to Business (B2B) Menggunakan Metode Root Cause Analysis (RCA) (Studi Kasus PT. X, Periode September 2019 -Agustus 2020).
- Lubis, Adyanata. 2016. Basis Data Dasar Untuk Mahasiswa Ilmu Komputer. Yogyakarta: CV. Budi Utama.
- Muharto, M., Hasan, S. and Ambarita, A., Penggunaan Model E-learning dalam Meningkatkan Hasil Belajar Mahasiswa pada Materi Microprocessor. Indonesian Journal on Information Systems, 2(1).
- Mulyani, S., 2017. Metode Analisis dan Perancangan Sistem. Abdi Sistematika.
- Mulyani, Sri. 2016. Sistem Informasi Manajemen. Bandung. Abdi Sistematika. Jeperson Hutahaeen. 2015. Konsep Sistem Informasi. Deepublish. Yogyakarta.
- Muslihudin, Muhammad dan Oktavianto. “Analisis dan Perancangan Sistem Informasi Menggunakan Model Terstruktur dan UML”. Yogyakarta: CV. Andi Offset, 2016.
- Ni Ketut Dewi Ari Jayanti, S.T., M. K., & Kadek Sumiari, S. Kom., M. M. S. . (2018). TEORI BASIS DATA (1st ed.; E. Risanto, ed.). Yogyakarta: Penerbit Andi.
- Nica Cahyani, Djamar T. F Lumban Batu, Sulistiono. 2016. Kandungan Logam Berat Pb, Hg, Cd, Dan Cu pada Daging Ikan Rejung (Sillago Sihama) di Estuari Sungai Donan, Cilacap, Jawa Tengah. JPHPI 2016, Volume 19 Nomor 3.
- Noki Ardian Madyaningrum, Sulastri. 2019. “Analisa Prediksi Kekambuhan Kanker Payudara Dengan Menggunakan Algoritma K-Nearest Neighbor”. Proceeding SINTAK 2019 ISBN: 978-602-8557-20-7.
- Nur Ghaniaviyanto Ramadhan dan Faisal Dharma Adhinata. 2021. “Teknik SMOTE dan Gini Score dalam Klasifikasi Kanker Payudara”. STITEK Bina Taruna Gorontalo Vol. 9 No. 2, Desember 2021, Hal. 125-134, ISSN: 2337-4101 E-ISSN: 2686-553X.
- Nurhadi, Slamet, dkk. 2019. Sistem Basis Data: Mysql di Ubuntu. Jakarta : Yayasan Manajemen dan Teknologi Nasional.
- Pamungkas, C. A., 2017. Pengantar Dan Implementasi Basis Data. 1 penyunt. Yogyakarta: Deepublish.
- Panny Agustia Rahayuningsih. 2019. “Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Tingkat Kematian Dini Kanker Dengan Dataset Early Death Cancer”. Jurnal Teknik Informatika Kaputama (JTIK) Vol. 3 , No. 2, Juli 2019, ISSN: 2548-9704.
- Risdiansyah, D. (2017). Perancangan Sistem Informasi Bimbingan Konseling Berbasis Desktop pada SMA Kemala Bhayangkari 1 Kubu Raya, V(2), 86- 91.
- Subandi & Syahidi, Aulia Akhriann 2018. Basis Data: Teori Dan Praktik Menggunakan Microsoft Office Access. Banjarmasin: Poliban Press.
- Sulianta, Feri. 2017. Teknik Perancangan Arsitektur Sistem Informasi. Andi: Yogyakarta.
- Umi Salamah dan Fata Nidaul Khasanah. 2017. Pengujian Sistem Informasi Penjualan Undangan Pernikahan Online Berbasis Web Menggunakan Black Box Testing. Information Management For Educators And Professionals Vol.2, No. 1, Desember2017, 35-46E-ISSN: 2548-3331.
- Wibowo, A. T., Saikhu, A., & Soelaiman, R. (2016). Implementasi Algoritma Deteksi SPAM yang Tersisipi Informasi Citra dengan Metode SVM dan Random Forest.



- Yayat Rukayat. 2017. "Kualitas Pelayanan Publik Bidang Administrasi Kependudukan di Kecamatan Pasirjambu". Jurnal Ilmiah Magister Ilmu Administrasi (JIMIA) No.2 Tahun XI Juni Vol 11, No 2 2017.
- Zufria, Ilka dan M. Hasan Azhari. 2017. Web-Based Applications in Calculation of Family Heritage (Science of Faroidh). Jurnal Sistem Informasi Vol.1 No.1: 2579-5341.