

SKRIPSI
KLASIFIKASI DALAM MENDETEKSI PENYAKIT KANKER
PAYUDARA MENGGUNAKAN NAÏVE BAYES



NOVAYANTI KLAU

NIM 205410095

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2024

SKRIPSI

KLASIFIKASI DALAM MENDETEKSI PENYAKIT KANKER

PAYUDARA MENGGUNAKAN NAIVE BAYES

Diajukan sebagai salah satu syarat untuk menyelesaikan studi

Program sarjana

Program Studi Informatika

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

Disusun Oleh

NOVAYANTI KLAU

NIM 205410095

PROGRAM STUDI INFORMATIKA

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFOMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2024

HALAMAN PERSETUJUAN

UJIAN SKRIPSI

Judul : **Klasifikasi dalam Mendeteksi Penyakit Kanker
Payudara Menggunakan Naïve Bayes**

Nama : **Novayanti klau**

NIM : **205410095**

Jurusan : **Informatika**

Semester : **Genap**

Tahun Akademik : **2023/2024**



Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Skripsi.

Yogyakarta, 02 Agustus 2024

Dosen pembimbing,

M. Agung Nugroho, S.Kom.,M.Kom

NIDN. 0507078501

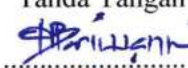


HALAMAN PENGESAHAN
SKRIPSI
KLASIFIKASI DALAM MENDETEKSI PENYAKIT KANKER
PAYUDARA MENGGUNAKAN NAIVE BAYES

Telah dipertahankan didepan Dewan Penguji Skripsi dan dinyatakan
diterima untuk memenuhi sebagai persyaratan guna memperoleh Gelar

Sarjana Komputer
Program Studi Informatika
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia
Yogyakarta

Yogyakarta, 02 Agustus 2024

Mengesahkan,

Dewan Penguji	NIDN	Tanda Tangan
1. Sari Iswanti, S.Si., M.Kom.	0508027202	
2. Ariesta Damayanti, S.Kom., M.Cs.	0020047801	
3. M. Agung Nugroho, S.Kom., M.Kom	0507078501	

Mengetahui

Ketua Program Studi Informatika


Dini Fakta Sari, S.T., M.T.

NPP : 121172

PERNYATAAN KEASLIAN SKRIPSI

Dengan ini penulis menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar sarjana komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan penulis tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 02 Agustus 2024



Novayanti Klau

NIM : 205410095

HALAMAN PERSEMBAHAN

Puji Syukur penulis panjatkan kepada Tuhan Yesus dan Bunda Maria, yang telah memberikan anugerah kesehatan dan kesanggupan, sehingga penulis masih diberikan kesempatan untuk menyelesaikan skripsi ini sebagai salah satu syarat untuk mendapatkan gelar sarjana, karna penulis percaya bahwa tanpa campur tangan dari mereka penulis tidak akan menyelesaikan skripsi ini. Dengan terselesainya tugas akhir ini penulis ingin mempersembahkannya kepada orang-orang yang penulis cintai dan sayangi, yakni:

1. Tuhan Yesus Kristus dan Bunda Maria yang begitu sangat baik memberikan hikmat, kesanggupan, penopang dan kesehatan sehingga penulis dapat menyelesaikan skripsi ini
2. Keluarga tercinta, Bapa Martinus, ibu Kristina, dan ibu Elisabeth, kk Elsa dan adik-adik saya yang telah memberikan semangat, dukungan dan doa dan menjadi tempat berkeluh kesah tetapi juga menjadi sumber kekuatan dan semangat dalam menyelesaikan skripsi ini
3. Teman-teman Informatika, dan teman-teman terdekat saya Angel, Nur, Hawila yang telah membantu dan memberikan semangat dalam mengerjakan skripsi dan selama perkuliahan.
4. Kepada Pemilki Tanggal 07-09-19 dan Sabahat penulis Selviani yang memberikan dukungan dan semangat dan menjadi tempat keluh kesah penulis.
5. Kepada diri sendiri terima kasih telah bertahan dan berjuang untuk sampai sejauh ini dan menyelesaikan skripsi ini, terima kasih juga

tidak menyerah menjadi kuat dan terima kasih sudah mencintai diri ini
dengan sangat baik.

MOTTO

“Yakin dan Percaya Bahwa semua Perkara dapat dilalui bersama dengan Tuhan”.

-Novayanti Klau-

“Direndahkan dimata Manusia, ditinggikan dimata Tuhan”

“Aku ditolak dengan hebat sampai jatuh, tetapi Tuhan menolong aku”

(Mazmur 118:13)

“Aku tahu, bahwa Engkau sanggup melakukan segala sesuatu dan tidak ada
rencana-Mu yang gagal”

(Ayub 42:2)

“Jangan takut, percaya saja

(Markus 5:36)

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yesus dan Bunda Maria, atas Anugerah-Nya baik berupa sehat fisik maupun akal pikiran yang senantiasa menjadi penopang, mendampingi penulis dan menjadi teman terbaik penulis sepanjang hidupnya, sehingga penulis dapat menyelesaikan Skripsi ini meskipun jauh dari kata sempurna.

Penyusunan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan ucapan terima kasih Kepada:

1. Ibu Sri Redjeki, S.Si., M.Kom., Ph.D. selaku Rektor Universitas Teknologi Digital Indonesia.
2. Bapak Dr. Bambang Purnomosidi DP, S. E. Akt., S. Kom., MMSI selaku Dekan Fakultas Teknologi Informasi.
3. Ibu Dini Fakta Sari, S.T., M.T. dan Ibu Femi Dwi Astuti, S.Kom., M.Cs., Ketua dan Sekretaris Program Studi Informatika Universitas Teknologi Digital Indonesia.
4. Bapak M. Agung Nugroho, M.Kom selaku dosen pembimbing skripsi yang telah meluangkan waktunya, dan memberikan bimbingan selama proses pengerjaan skripsi.
5. Ibu Sari Iswanti, S.Si., M.Kom. dan Ibu Ariesta Damayanti, S.Kom.,M.Cs. selaku dosen penguji pada sidang skripsi
6. Keluarga Tercinta, Terutama Bapak Martinus mama Kristina, dan mama Elisabeth yang selalu mendukung, memberikan semangat dan mendoakan penulis agar menyelesaikan skripsi ini.

7. Kepada pemilik tanggal 07-09-2018, sahabat penulis Selviani, dan Teman-teman mahasiswa Universitas Teknologi Digital Indonesia yang tidak bisa penulis sebutkan satu persatu yang telah memberikan dukungan dan terlibat dalam membantu terselesainya skripsi ini.

Menyadari akan kekurangan dan kesalahan penulis dalam penyusunan skripsi ini karena keterbatasan ilmu penulis, maka penulis mengharapkan masukan berupa saran yang membangun dari semua pihak.

Semoga skripsi ini dapat memberikan manfaat, baik penulis sendiri maupun bagi masyarakat luas.

Yogyakarta, 02 Agustus 2024



Novayanti Klau

205410095

DAFTAR ISI

SKRIPSI.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xii
INTISARI.....	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah	2
1.3 Ruang Lingkup.....	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	14
2.1 Tinjauan Pustaka	14
2.1 Dasar Teori.....	18
2.2.1 Kanker Payudara	18
2.2.2 Klasifikasi	19
2.2.3 Naïve Bayes	19
2.2.4 Preprocessing Data.....	20
2.2.5 Confusion Matrix	20
2.2.6 Python	22
2.2.7 Streamlit	22
2.2.8 Alur Penelitian	22
BAB III	24
3.1 Bahan/Data.....	24

3.2 Analisis Kebutuhan	26
3.2.1 Kebutuhan Perangkat Keras	26
3.2.2 Kebutuhan perangkat lunak.....	26
3.2.3 Kebutuhan Masukan (input).....	27
3.2.4 Kebutuhan Proses.....	27
3.2.5 Kebutuhan output.....	28
3.3 Analisis dan Perancangan sistem	28
3.3.1 <i>Usecase</i> Diagram	28
3.3.2 Perancangan Antarmuka	29
1. Halaman Login.....	29
2. Halaman Kanker payudara.....	30
3. Halaman Input Data	31
2. Halaman Hasil Klasifikasi Data	32
BAB IV	22
4.1 Implementasi	22
4.1.1 Import Dataset.....	22
4.1.2 Preprocessing Data.....	22
4.1.3 Membagi Data <i>Training</i> dan Data <i>Testing</i>	24
4.1.4 Implementasi <i>Naive Bayes</i>	24
4.1.5 Hasil Akurasi.....	25
4.2 Pengujian dan pembahasan Aplikasi <i>Web</i>	27
4.2.1 Antarmuka Aplikasi <i>web</i>	29
4.2.2 Hasil Uji Coba.....	31
BAB V.....	31
5.1 Kesimpulan	31
5.2 Saran.....	31
DAFTAR PUSTAKA	33
LAMPIRAN CARA MENJALANKAN PROGRAM.....	34
LISTING PROGRAM	35
LAMPIRAN PENDADARAN	40
SURAT KETERANGAN PERSETUJUAN PUBLIKASI.....	42

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Penelitian	23
Gambar 3. 1 Alur use case diagram	28
Gambar 3. 2 Halaman login	29
Gambar 3. 3 Halaman Home.....	30
Gambar 3. 4 Halaman input	31
Gambar 3. 5 Hasil data klasifikasi	32
Gambar 4. 1 Read Dataset.....	22
Gambar 4. 2 Memeriksa Missing Value	23
Gambar 4. 3 pemeriksaan missing value.....	23
Gambar 4. 4 Memeriksa duplikat data	23
Gambar 4. 5 memeriksa duplikat data.....	23
Gambar 4. 6 Pembagian data Training dan data Testing	24
Gambar 4. 7 Implementasi naive bayes	25
Gambar 4. 8 tampilan confusion matrix.....	25
Gambar 4. 9 Hasil Akurasi.....	25
Gambar 4. 10 Hasil Akurasi.....	26
Gambar 4. 11 Halaman Login.....	29
Gambar 4. 12 Halaman Home.....	30
Gambar 4. 13 Halaman Input data	31
Gambar 4. 14 Uji Coba Pasien Kanker Payudara	32
Gambar 4. 15 Uji Coba Pasien control sehat	33
Gambar 4. 16 Halaman hasil klasifikasi data.....	33

DAFTAR TABEL

Tabel 2. 1 Perbandingan.....	14
Tabel 2. 2 Penjelasan Confusion Matrix	20
Tabel 3. 1 Dataset Kanker Payudara	24
Tabel 4. 1 Hasil perhitungan confusion matrix	26
Tabel 4. 2 Pengujian Blackbox Sistem	27

INTISARI

Kanker payudara merupakan kanker paling umum dan paling banyak terjadi pada wanita di seluruh dunia. Kanker payudara menempati urutan pertama dengan jumlah kasus terbanyak di Indonesia, dan menjadi salah satu penyumbang kematian pertama akibat kanker payudara. Menurut data statistik Globocan tahun 2020, tercatat jumlah kasus baru kanker payudara mencapai 68.858 (16,6%) dari total data 396.914 di Indonesia. Kemudian, untuk jumlah kematiannya mencapai lebih dari 22 ribu jiwa kasus. Tingginya angka kematian karena kanker payudara di Indonesia perlu dicermati dengan tindakan pencegahan dan deteksi dini.

Klasifikasi adalah pengelompokan data ke dalam kategori yang telah ditentukan sebelumnya. Dalam klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasi penyakit kanker payudara berdasarkan *dataset Coimbra UCI Machine Learning Repository* yang diperoleh dari *kaggle* dengan metode *naïve bayes*. Metode *naïve bayes* merupakan Algoritma klasifikasi yang dapat dikatakan mampu dalam melakukan klasifikasi data dengan baik, berdasarkan kejadian sebelumnya. Hasil pengujian klasifikasi *naïve bayes* pada dataset sebanyak 116 data menggunakan *Confusion Matrix* pada data *training* 80% dan *testing* 20% diperoleh akurasi sebesar 80%, Precision 100% dan recall 67%.

Kata kunci: Kanker Payudara, klasifikasi, Naive Bayes

ABSTRACT

Breast cancer is the most common and most prevalent cancer in women worldwide. Breast cancer ranks first with the highest number of cases in Indonesia, and is one of the first contributors to death from breast cancer. According to Globocan statistics in 2020, the number of new cases of breast cancer reached 68,858 (16.6%) of the total data of 396,914 in Indonesia. Then, the number of deaths reached more than 22 thousand cases. The high mortality rate due to breast cancer in Indonesia needs to be observed with preventive measures and early detection.

Classification is the grouping of data into predetermined categories. Classification involves the process of examining the characteristics of the object and putting the object into one of the predefined classes.

This study aims to identify and classify breast cancer diseases based on the Coimbra UCI Machine Learning Repository dataset obtained from kaggle data with the naïve bayes method. The naïve bayes method is a classification algorithm that can be said to be able to classify data well, based on previous events. The results of naïve bayes classification test results on a datasets of 116 data using the Confusion Matrix on 80% training data and 20% testing obtained an accuracy of 80%, Precision 100% and recall 67%.

Keywords: *breast cancer, Classification, Naive Bayes*