

**PENGLASIFIKASIAN KANKER SERVIKS
MENGUNAKAN METODE *FUZZY NAÏVE BAYES* DENGAN
*BOOTSTRAP SAMPLING***

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains Bidang Studi Matematika**

Oleh :

KHOIROTUN NISA

08011281924033



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2023

LEMBAR PENGESAHAN

**PENGLASIFIKASIAN KANKER SERVIKS
MENGUNAKAN METODE *FUZZY NAÏVE BAYES* DENGAN
*BOOTSTRAP SAMPLING***

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains Bidang Studi Matematika**

Oleh

KHOIROTUN NISA

08011281924033

Pembimbing Kedua



Endang Sri K, S.Si., M.Si
NIP. 197702082002122003

Indralaya, April 2023

Pembimbing Utama



Dr. Yulia Resti, S.Si., M.Si
NIP. 197307191997022001

Mengetahui,

Ketua Jurusan Matematika



Drs. Sugandi Yahdin, M.M
NIP. 195807271986031003

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Khoirotun Nisa
NIM : 08011281924033
Fakultas/Jurusan : MIPA/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan starata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat dalam skripsi ini berasal dari penulis lain baik yang dipublikasikan atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 3 April 2023



Penulis

LEMBAR PERSEMBAHAN

MOTTO

“Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya”

(Q.S Al-Baqarah: 286)

“I feel like the possibility of all those possible is just another possibility that can possibly happen”

(Mark lee)

Skripsi ini kupersembahkan kepada:

- 1. Allah SWT**
- 2. Kedua Orangtua**
- 3. Saudaraku**
- 4. Keluarga Besarku**
- 5. Semua Dosen dan Guruku**
- 6. Sahabat dan temanku**
- 7. Almamaterku**

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur kehadiran Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Pengklasifikasian Kanker Serviks Menggunakan Metode *Fuzzy Naïve Bayes* dengan *Bootstrap Sampling*”** dengan baik. Shalawat beserta salam semoga selalu tercurahkan kepada baginda kita Nabi Muhammad SAW beserta keluarga, sahabat dan para pengikutnya hingga akhir zaman. Skripsi ini merupakan salah satu persyaratan yang harus dipenuhi penulis untuk meraih gelar Sarjana Sains bidang Studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Dengan segala hormat dan kerendahan hati penulis mengucapkan terima kasih yang sebesar-besarnya untuk kedua orang tua tercinta, yaitu **Bapak Danmar** dan **Ibu Huzaima** yang telah mendidik, menuntun, menasehati, memberi semangat dan perhatian serta selalu berdo'a yang terbaik untuk anaknya.

Penulis juga mengucapkan terima kasih kepada Ibu **Dr. Yulia Resti, M.Si** selaku Dosen Pembimbing Utama sekaligus Dosen Pembimbing Akademik dan Ibu **Endang Sri Kresnawati, M.Si** selaku Dosen Pembimbing Kedua yang telah bersedia meluangkan waktu, tenaga, pikiran untuk bimbingan dan pengarahan penuh perhatian, pengertian dan kesabaran sehingga skripsi ini dapat diselesaikan.

Penulis juga mengucapkan terima kasih kepada:

1. Bapak **Drs. Sugandi Yahdin, M.M** selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya sekaligus

Dosen Pembahas I yang telah memberikan waktu serta saran dan tanggapannya yang bermanfaat bagi penulis dalam mengerjakan skripsi.

2. Ibu **Dr. Dian Cahyawati Sukanda, M.Si** selaku Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya

3. Ibu **Irmeilyana, M.Si** selaku Dosen Pembahas II yang telah memberikan tanggapan, kritik, dan saran yang sangat bermanfaat untuk perbaikan dan penyelesaian skripsi ini.

4. Bapak **Drs. Robinson Sitepu, M.Si** selaku Ketua Seminar yang telah memberikan arahan, saran, serta mengatur jalannya seminar sehingga dapat berjalan dengan baik.

5. Ibu **Des Alwine Zayanti, M.Si** selaku Sekretaris Seminar yang telah membantu dan memberikan catatan serta masukan yang sangat berguna bagi penulis.

6. **Seluruh Dosen di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam** yang telah memberikan ilmu yang bermanfaat, nasihat serta motivasi selama penulis menjalani perkuliahan.

7. Bapak **Irwansyah** dan Ibu **Hamidah** selaku Pegawai tata usaha Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah membantu penulis selama perkuliahan.

8. Keluarga Besarku terutama nenekku **Hj. Zubaida**, Kakak dan adikku **Randa Sholihin** dan **Delima** terima kasih untuk kasih sayang, motivasi, dukungan, perhatian dan do'a yang selalu dipanjatkan selama ini untuk keberhasilanku.

9. Sahabat-sahabatku, **Mentari, Aulia, Chika, Laurit, Sinta** dan seluruh teman-teman angkatan 2019 terimakasih atas dukungan, semangat dan kebersamaannya.
10. Teman seperjuangan skripsi **Leliani, Ayu, Tiara, Fitriana, Elisa, Hutvina,** dan **Uci**, terima kasih untuk saling menguatkan, berbagi ilmu dan saling mengajarkan dan bantuan selama proses penyelesaian skripsi ini.
11. Teman kamarku **Aprilia Sari**, terima kasih telah menemani dan mendengarkan keluh kesah selama proses penyelesaian skripsi ini.
12. Kakak tingkat 2018 kak **Nurafni**, kak **Adin** terima kasih atas arahan dan bantuan dalam proses skripsi, dan seluruh kakak tingkat angkatan 2017, 2018, serta adik tingkat 2020 dan 2021 atas bantuan selama perkuliahan.
13. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah memeberikan bantuan dalam menyelesaikan skripsi ini terima kasih banyak. Semoga segala kebaikan yang diberikan mendapatkan balasan dari Allah Subhanahu Wa Ta'ala.

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam dan semua pihak yang memerlukan.

Wassalamu'alaikum warahmatullahi wabarakatuh.

Indralaya, April 2023

Penulis

**CLASSIFICATION OF CERVICAL CANCER
USING FUZZY NAÏVE BAYES METHOD WITH
BOOTSTRAP SAMPLING**

By:

**KHOIROTUN NISA
08011281924033**

ABSTRACT

Cervical cancer ranks second in the top cause of death in women in various countries, including Indonesia. The Global Burden Of Cancer Study notes that in 2020 there will be 604,127 new cases of cervical cancer worldwide. Around 351,720 new cases of cervical cancer on the Asian continent, with 190,874 cases of cervical cancer occurring in Southeast Asia. The high cases of cervical cancer are related to the late examination and diagnosis of cervical cancer. Therefore, a study is needed that discusses the classification for diagnosing cervical cancer patients with high accuracy. The purpose of this study was to classify cervical cancer based on the results of pap smear cell image extraction using the naïve bayes and fuzzy naïve bayes methods with bootstrap sampling. The data used in this study is a dataset of pap smear images of cervical cancer grade 2 and class 7 from Herlev University Hospital. The results of this study indicate that the classification of cervical cancer using the naïve bayes fuzzy method is better than naïve bayes, both for 2 classes and 7 classes. Classification using the naïve bayes method for 2 classes with bootstrap sampling produces an average value of accuracy, precision, recall and specificity of 89.27%, 86.53%, 67.83% and 96.42%, respectively. While the classification using the naïve bayes method for 7 classes with bootstrap sampling produces an average accuracy, precision, recall and specificity of 82.12%, 40.10%, 45.34% and 87.30%, respectively. Classification using the fuzzy naïve bayes method for 2 classes with bootstrap sampling produces an average value of accuracy, precision, recall and specificity of 90.26%, 89.93%, 61.94% and 97.54%, respectively. While the classification using the fuzzy naïve bayes method for 7 classes with bootstrap sampling produces an average accuracy, precision, recall and specificity of 86.94%, 60.2%, 55.17% and 90.51%, respectively.

Keywords: Cervical Cancer, Fuzzy Set, Naïve Bayes, Fuzzy Naïve Bayes, Bootstrap Sampling

**PENGLASIFIKASIAN KANKER SERVIKS
MENGUNAKAN METODE *FUZZY NAÏVE BAYES* DENGAN
*BOOTSTRAP SAMPLING***

Oleh:

**KHOIROTUN NISA
08011281924033**

ABSTRAK

Kanker serviks menempati peringkat kedua teratas penyebab kematian pada perempuan di berbagai negara termasuk Indonesia. *Global Burden Of Cancer Study* mencatat bahwa pada tahun 2020 terdapat kasus baru kanker serviks sebanyak 604.127 kasus di seluruh dunia. Sekitar 351.720 kasus baru kanker serviks di benua Asia, dengan 190.874 kasus kanker serviks yang terjadi di Asia Tenggara. Tingginya kasus kanker serviks ini berkaitan dengan terlambatnya pemeriksaan dan diagnosis penyakit kanker serviks. Oleh karena itu, diperlukan suatu penelitian yang membahas tentang klasifikasi untuk mendiagnosis pasien penderita kanker serviks dengan akurasi yang tinggi. Tujuan penelitian ini yaitu melakukan klasifikasi kanker serviks berdasarkan hasil ekstraksi citra sel *pap smear* menggunakan metode *naïve bayes* dan *fuzzy naïve bayes* dengan *bootstrap sampling*. Data yang digunakan dalam penelitian ini adalah dataset hasil citra sel *pap smear* kanker serviks 2 kelas dan 7 kelas yang berasal dari Herlev University Hospital. Hasil penelitian ini menunjukkan bahwa pengklasifikasian kanker serviks menggunakan metode *fuzzy naïve bayes* lebih baik dibandingkan dengan *naïve bayes*, baik itu untuk 2 kelas maupun 7 kelas. Klasifikasi menggunakan metode *naïve bayes* untuk 2 kelas dengan *bootstrap sampling* menghasilkan nilai rata-rata *accuracy*, *precision*, *recall* dan *specificity* masing-masing sebesar 89,27%, 86,53%, 67,83% dan 96,42%. Sedangkan klasifikasi menggunakan metode *naïve bayes* untuk 7 kelas dengan *bootstrap sampling* menghasilkan rata-rata *accuracy*, *precision*, *recall* dan *specificity* masing-masing sebesar 82,12%, 40,10%, 45,34% dan 87,30%. Klasifikasi menggunakan metode *fuzzy naïve bayes* untuk 2 kelas dengan *bootstrap sampling* menghasilkan nilai rata-rata *accuracy*, *precision*, *recall* dan *specificity* masing-masing sebesar 90,26%, 89,93%, 61,94% dan 97,54%. Sedangkan klasifikasi menggunakan metode *fuzzy naïve bayes* untuk 7 kelas dengan *bootstrap sampling* menghasilkan rata-rata *accuracy*, *precision*, *recall* dan *specificity* masing-masing sebesar 86,94%, 60,2%, 55,17% dan 90,51%.

Kata Kunci: Kanker Serviks, Himpunan *fuzzy*, *Naïve Bayes*, *Fuzzy Naïve Bayes*, *Bootstrap Sampling*

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSEMBAHAN	iii
KATA PENGANTAR	iv
ABSTRACT	vii
ABSTRAK	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	6
1.5 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Kanker Serviks	7
2.2 <i>Statistical Machine Learning</i>	8
2.3 Peluang	8
2.4 Klasifikasi	9
2.5 Diskritisasi Data	10
2.6 Bootstrap Sampling	11
2.7 Metode <i>Naïve Bayes</i>	12
2.8 <i>Laplace Smoothing</i>	15
2.9 Himpunan <i>Fuzzy</i>	15
2.10 Fungsi Keanggotaan	16
2.11 Fuzzy <i>Naïve Bayes</i>	20
2.12 <i>Confusion Matrix</i>	21
BAB III METODOLOGI PENELITIAN	24

3.1	Tempat Penelitian.....	24
3.2	Waktu Penelitian	24
3.3	Data Penelitian	24
3.4	Metode Penelitian.....	27
BAB IV HASIL DAN PEMBAHASAN		30
4.1	Deskripsi Data	30
4.2	Klasifikasi Menggunakan Metode <i>Naïve Bayes</i>	32
4.2.1	Diskritisasi Data.....	32
4.2.2	Partisi Data.....	35
4.2.3	<i>Bootstrap Sampling</i>	36
4.2.4	<i>Naïve Bayes</i> untuk 2 kelas	37
4.2.5	<i>Naïve Bayes</i> untuk 7 kelas	42
4.3	Klasifikasi Menggunakan Metode <i>Fuzzy Naïve Bayes</i>	48
4.3.1	Menentukan Himpunan Universal	49
4.3.2	Menentukan Nilai Keanggotaan <i>Fuzzy</i>	56
4.3.3	Partisi Data.....	59
4.3.4	<i>Bootstrap Sampling</i>	61
4.3.5	<i>Fuzzy Naïve Bayes</i> untuk 2 kelas	61
4.3.6	<i>Fuzzy Naïve Bayes</i> untuk 7 kelas	67
4.4	Analisis Hasil.....	73
BAB V KESIMPULAN DAN SARAN		76
5.1	Kesimpulan.....	76
5.2	Saran	76
DAFTAR PUSTAKA		78
LAMPIRAN.....		82

DAFTAR TABEL

Tabel 2.1 <i>Confusion matrix</i>	21
Tabel 2.2 <i>Multiclass confusion matrix</i>	22
Tabel 3.1 Karakteristik sel <i>pap smear</i>	25
Tabel 3.2 Variabel Prediktor	26
Tabel 3.3 Lanjutan Variabel Prediktor	27
Tabel 4.1 <i>Dataset</i> sel <i>pap smear</i> untuk 2 kelas.....	30
Tabel 4.2 <i>Dataset</i> sel <i>pap smear</i> untuk 7 kelas.....	30
Tabel 4.3 Nilai minimum, maksimum dan range setiap varaibel prediktor.....	33
Tabel 4.4 Hasil diskritisasi data untuk 2 kelas	34
Tabel 4.5 Hasil diskritisasi data untuk 7 kelas	34
Tabel 4.6 Data latih untuk 2 kelas metode NB	35
Tabel 4.7 Data uji untuk 2 kelas metode NB	35
Tabel 4.8 Data latih untuk 7 kelas metode NB	35
Tabel 4.9 Data uji untuk 7 kelas metode NB	36
Tabel 4.10 Data latih untuk 2 kelas hasil <i>Bootstrap Sampling</i> ke-12 metode NB	36
Tabel 4.11 Data latih untuk 7 kelas hasil <i>Bootstrap Sampling</i> ke-12 metode NB	36
Tabel 4.12 Nilai Peluang <i>Likelihood</i> untuk 2 kelas	39
Tabel 4.13 Hasil Klasifikasi metode <i>Naïve Bayes</i> untuk 2 kelas.....	41
Tabel 4.14 <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	41
Tabel 4.15 Jumlah data setiap kategori variabel prediktor terhadap variabel respon	43
Tabel 4.16 Nilai Peluang <i>Likelihood</i> untuk 7 kelas	44
Tabel 4.17 Hasil Klasifikasi metode <i>Naïve Bayes</i> untuk 7 kelas.....	47
Tabel 4.18 Tabel klasifikasi metode Metode <i>Naïve Bayes</i>	47
Tabel 4.19 <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	48
Tabel 4.20 Interval himpunan <i>fuzzy</i> variabel prediktor X_1	49
Tabel 4.21 Himpunan <i>fuzzy</i> variabel prediktor X_1	50
Tabel 4.22 Interval himpunan <i>fuzzy</i> variabel prediktor X_2	52
Tabel 4.23 Himpunan <i>fuzzy</i> variabel prediktor X_2	52
Tabel 4.24 Interval himpunan <i>fuzzy</i> variabel prediktor X_3	54
Tabel 4.25 Himpunan <i>fuzzy</i> variabel prediktor X_3	54
Tabel 4.26 Nilai keanggotaan terbesar variabel prediktor X_1	57
Tabel 4.27 Nilai keanggotaan terbesar variabel prediktor X_2	58
Tabel 4.28 Nilai keanggotaan terbesar variabel prediktor X_3	59
Tabel 4.29 Nilai keanggotaan terbesar semua variabel prediktor	59
Tabel 4.30 Data latih untuk 2 kelas metode FNB	60

Tabel 4.31 Data uji untuk 2 kelas metode FNB	60
Tabel 4.32 Data latih untuk 7 kelas metode FNB	60
Tabel 4.33 Data uji untuk 7 kelas metode FNB	60
Tabel 4.34 Data latih untuk 2 kelas hasil <i>Bootstrap Sampling</i> ke-12 metode FNB	61
Tabel 4.35 Data latih untuk 7 kelas hasil <i>Bootstrap Sampling</i> ke-12 metode FNB	61
Tabel 4.36 Nilai Peluang <i>Likelihood</i> untuk 2 kelas	62
Tabel 4.37 Hasil Klasifikasi metode <i>Fuzzy Naïve Bayes</i> untuk 2 kelas.....	66
Tabel 4.38 <i>Confusion Matrix</i> Metode <i>Fuzzy Naïve Bayes</i>	66
Tabel 4.39 Jumlah data setiap kategori variabel prediktor terhadap variabel respon	68
Tabel 4.40 Nilai Peluang <i>Likelihood</i> untuk 7 kelas	69
Tabel 4.41 Hasil Klasifikasi metode <i>Fuzzy Naïve Bayes</i> untuk 7 kelas.....	72
Tabel 4.42 Tabel klasifikasi metode Metode <i>Fuzzy Naïve Bayes</i>	72
Tabel 4.43 <i>Confusion Matrix</i> Metode <i>Fuzzy Naïve Bayes</i>	73
Tabel 4.44 Perbandingan hasil Metode <i>Niave Bayes</i> dan <i>Fuzzy Naïve Bayes</i>	74

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Tahapan <i>Bootstrap Sampling</i>	12
Gambar 2.2 Kurva Linier Naik	17
Gambar 2.3 Kurva Linier Turun	18
Gambar 2.4 Kurva Segitiga.....	19
Gambar 4.1 Sebaran frekuensi 2 kelas kanker serviks.....	31
Gambar 4.2 Sebaran frekuensi 7 kelas kanker serviks.....	32

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker merupakan salah satu penyebab utama kematian terbanyak di seluruh dunia. Kanker mempunyai berbagai macam jenis salah satunya kanker serviks. Kanker serviks menempati peringkat kedua teratas penyebab kematian pada perempuan di berbagai negara termasuk Indonesia (Wahyuni & Syswianti, 2018). Berdasarkan *Global Burden Of Cancer Study* (2020) mencatat bahwa terdapat kasus baru kanker serviks sebanyak 604.127 kasus di seluruh dunia. Sekitar 351.720 kasus baru kanker serviks di benua Asia, dengan 190.874 kasus kanker serviks yang terjadi di Asia Tenggara (Sung et al., 2021). Adapun di Indonesia sendiri pada tahun 2020 terdapat kasus baru kanker serviks sebanyak 36.633 kasus dan berkisaran 21.003 jiwa meninggal. Pada tahun 2019 terdapat kasus kanker serviks sebanyak 23,4 per 100.000 penduduk dengan rata-rata kematian 13,9 per 100.000 penduduk (Naufaldi et al., 2020).

Tingginya kasus kanker serviks ini penyebab utamanya adalah infeksi *Human Papiloma Virus* (HPV) (Akbar & Sandfreni, 2021). Pada prosesnya kanker serviks cenderung penyakit yang diam pada tahap prakanker dan kanker awal serta tidak menimbulkan gejala atau keluhan yang diterima oleh penderita (Mastutik et al., 2015). Adanya hal ini menyebabkan sering kali kanker serviks tidak teridentifikasi sampai keadaan sudah tahap lanjut. Permasalahan tersebut juga berkaitan dengan terlambatnya pemeriksaan dan diagnosis penyakit kanker serviks (Wahyuni & Syswianti, 2018). Maka dari itu, perlu dilakukannya pemeriksaan awal bagi

wanita untuk mengantisipasi kemungkinan adanya sel yang abnormal terkait kanker serviks.

Untuk mendeteksi kanker serviks bisa dilakukan dengan tes *pap smear*, pemeriksaan HPV, dan pemeriksaan IVA yang didukung dengan tes pencitraan atau gambar medis seperti CT scan, mikroskopis dan MRI (Akbar & Sandfreni, 2021). Namun yang sering digunakan di Indonesia ada 2 yakni tes *pap smear* dan IVA. Tes *pap smear* dapat mendeteksi perubahan sel-sel rahim di bawah mikroskop untuk mengetahui apakah ada kemungkinan sel serviks akan menjadi kanker di masa yang akan datang. Menjalani tes *pap smear* merupakan langkah awal dalam mendeteksi sel abnormal sejak dini. Langkah ini bisa mengatasi dan menghentikan kemungkinan berkembangnya kanker serviks (Arifin, 2017). Namun untuk melakukan tes *pap smear* sangat membutuhkan waktu dan teknis sitologi yang terampil dalam membedakan sel-sel yang beragam (Wijaya et al., 2021). Adanya suatu sistem yang akurat yang bisa digunakan untuk mendiagnosis hasil *pap smear* akan sangat membantu bagi para sitologi. Maka dari itu, diperlukan suatu penelitian tentang klasifikasi dengan hasil *pap smear* dengan akurasi yang tinggi, dimana bisa menyamakan hasil diagnosis para sitologi.

Banyak metode yang bisa digunakan untuk melakukan klasifikasi data dengan hasil yang akurat dan baik, salah satunya yaitu metode *fuzzy naive bayes*. Metode *fuzzy naive bayes* adalah kombinasi himpunan *fuzzy* dan *naive bayes* dalam perhitungan di dalamnya. Himpunan *fuzzy* bekerja dengan menggunakan derajat keanggotaan dari sebuah nilai yang kemudian digunakan untuk menentukan hasil yang ingin dihasilkan berdasarkan atas spesifikasi yang telah

ditentukan. *Naive bayes* merupakan pengklasifikasian dengan memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya (Sastrawan et al., 2019). Klasifikasi *Bayes* didasarkan pada teorema *bayes*, diambil dari nama seorang ahli yaitu Thomas Bayes (Septiani et al., 2017). Teorema tersebut dikombinasikan dengan *naive* di mana diasumsikan kondisi antar atribut saling bebas. Jadi, *naïve bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya (Bustami, 2013). Ada penelitian terdahulu tentang *fuzzy naïve bayes* yakni Fuad et al. (2019) dengan akurasi sebesar 78.2% dari 81 data. Ada juga Tütüncü & Kayaalp (2015) dengan tingkat akurasi 81%. Pada proses klasifikasi naïve bayes sering terjadinya naik turun akurasi karena sampling yang tidak tepat. Maka untuk mengatasi hal ini diperlukan proses sampling.

Bootstrap adalah suatu metode yang dapat bekerja tanpa membutuhkan asumsi distribusi karena sampel asli digunakan sebagai populasi. *Bootstrap* juga digunakan untuk mengatasi permasalahan dalam statistika baik masalah data yang sedikit, data yang menyimpang dari asumsinya maupun data yang tidak memiliki asumsi dalam distribusinya. *Bootstrap* adalah suatu metode yang berbasis komputer yang sangat potensial untuk dipergunakan pada masalah keakurasian. Metode *bootstrap* dilakukan dengan mengambil sampel dari sampel asli dengan ukuran sama dengan sampel asli dan dilakukan dengan pengembalian (Sungkono, 2013). Terdapat beberapa penelitian sebelumnya yang menggunakan *bootstrap sampling* yakni Informasi et al. (2019) menyatakan *bootstrap* terbukti dapat meningkatkan *accuracy* dari pengklasifikasi sebesar 5.41 %, dari pengujian KNN

mendapatkan *accuracy* 91,89%. Ada juga Agus et al. (2017) dengan sample *bootstrapping* mempunyai nilai akurasi yang lebih baik yaitu 5.4% (dari 91.52% menjadi 96.87%).

Penelitian tentang klasifikasi citra pap smear kanker serviks dengan menggunakan dataset Herlev sebelumnya telah dilakukan oleh beberapa peneliti antara lain Sun et al. (2017) dalam penelitiannya menggunakan metode *naïve bayes* dengan akurasi sebesar 91.71%, C4.5 akurasi sebesar 92.15% dan random forest 92.58%. Ada juga penelitian lain Athinarayanan dan Srinath (2016) dengan menggunakan SVM akurasi 94%, Bora (2017) menggunakan metode ensambel tingkat akurasi 96.51% untuk 2 kelas dan 91.71% untuk 3 kelas, dan nada Chankong et al. (2014) menggunakan metode FCM dengan tingkat akurasi 93.78% untuk 7 kelas, 99.27% untuk 2 kelas (Shanthi et al., 2022).

Berdasarkan penelitian-penelitian sebelumnya, penulis tertarik untuk melakukan penelitian tentang hasil citra *pap smear* dari data Herlev University menggunakan metode *naïve bayes* dengan *bootstrap sampling* untuk klasifikasi kanker serviks. Dimana *bootstrap sampling* tersebut diharapkan dapat meningkatkan kinerja dari metode *fuzzy naïve bayes*. Hasil kinerja diukur berdasarkan tingkat ketepatan klasifikasi, dimana pada penelitian ini dibatasi pada *accuracy*, *precision*, *recall*, dan *specificity*.

1.2 Perumusan Masalah

Rumusan masalah pada penelitian ini sebagai berikut.

1. Bagaimana pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas?

2. Bagaimana pengklasifikasian kanker serviks menggunakan metode *fuzzy naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas?
3. Bagaimana tingkat ketepatan dalam pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas?
4. Bagaimana tingkat ketepatan dalam pengklasifikasian kanker serviks menggunakan metode *fuzzy naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas?
5. Bagaimana perbandingan tingkat ketepatan dalam pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dan *fuzzy naïve bayes* dengan *bootstrap sampling*.

1.3 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut.

1. Dataset yang digunakan dalam penelitian ini hasil citra sel *pap smear* kanker serviks dari Herlev University Hospital dengan jumlah data 917 citra.
2. Data dipartisi menjadi 80% (733 citra) data latih dan 20% (184 citra) data uji.
3. Fungsi keanggotaan *fuzzy* yang digunakan ada 3 yaitu kurva linier turun, kurva linier naik dan kurva segitiga.
4. Persentase ketepatan klasifikasi dibatasi oleh nilai *accuracy*, *precision*, *recall*, dan *specificity*.

1.4 Tujuan Penelitian

Tujuan pada penelitian ini sebagai berikut.

1. Menentukan pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dan *fuzzy naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas.
2. Menentukan tingkat ketepatan dalam pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dan *fuzzy naïve bayes* dengan *bootstrap sampling* untuk 2 kelas dan 7 kelas.
3. Menentukan perbandingan tingkat ketepatan dalam pengklasifikasian kanker serviks menggunakan metode *naïve bayes* dan *fuzzy naïve bayes* dengan *bootstrap sampling*

1.5 Manfaat Penelitian

Manfaat pada penelitian ini sebagai berikut.

1. Penelitian ini dapat membantu tenaga medis terkait informasi dalam melakukan klasifikasi diagnosis pada pasien kanker serviks.
2. Penelitian ini diharapkan dapat menambah pengetahuan dan wawasan mengenai pengklasifikasin kanker serviks menggunakan metode *fuzzy naïve bayes* dengan *bootstrap sampling*.
3. Sebagai bahan referensi untuk penelitian mengenai kanker serviks dan metode *fuzzy naïve bayes*.

DAFTAR PUSTAKA

- Agus, T., Adib, S. M., & Karomi, A. (2017). *Penerapan Metode Sample Bootstrapping untuk Meningkatkan Performa kNearest Neighbor pada Dataset Berdimensi Tinggi*. *XII*(1), 9–14.
- Akbar, H., & Sandfreni, S. (2021). Klasifikasi Kanker Serviks Menggunakan Model Convolutional Neural Network Alexnet. *JIKO (Jurnal Informatika Dan Komputer)*, *4*(1), 44–51. <https://doi.org/10.33387/jiko.v4i1.2606>
- Ananta, P., Putra, D., Purnawan, I. K. A., Purnami, D., & Putri, S. (2018). *Sistem Pakar Diagnosa Penyakit Mata dengan Fuzzy Logic dan Naïve Bayes*. *6*(1), 35–46.
- Arifin, T. (2017). Implementasi Algoritma PSO Dan Teknik Bagging Untuk Klasifikasi Sel Pap Smear. *Jurnal Informatika*, *4*(2), 155–162.
- Azis, H., Tangguh Admojo, F., & Susanti, E. (2020). Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah. *Techno.Com*, *19*(3), 286–294. <https://doi.org/10.33633/tc.v19i3.3646>
- Berrar, D. (2018). Bayes' theorem and naive bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1–3*(September), 403–412. <https://doi.org/10.1016/B978-0-12-809633-8.20473-1>
- Fauziningrum, M. Pd, E., & Encis Indah Suryaningsih, S.T., M. P. (2021). Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim Amni). *Angewandte Chemie International Edition*, *6*(11), 951–952., 5–24.
- Fuad, N., Studi, P., Informatika, T., & Lamongan, U. I. (2019). *Kombinasi Algoritma Fuzzy Naive Bayes Pada Pemilihan. September*.
- Hidayah, U. R., Cholissodin, I., & Adikara, P. P. (2019). Klasifikasi Penyakit Kanker Serviks dengan Extreme Learning Machine. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *3*(7), 6575–6582. <http://j-ptiik.ub.ac.id>
- Hidayatulloh, T., Herliana, A., & Arifin, T. (2016). Klasifikasi Sel Tunggal Pap Smear Berdasarkan Analisis Fitur Berbasis Naïve Bayes Classifier Dan Particle Swarm Optimization. ., *4*(2), 186–193.
- Informasi, J. T., Siswanto, E., Sarjana, P., Informatika, T., & Dian, U. (2019). *Metode Sample Bootstrapping Pada K-Nearest Neighbor*. *14*, 13–23.
- Istiqomah, I. (2016). Penerapan Teorema Binomial Untuk Menentukan Peluang

- Kejadian (Kasus :Percobaan Pelemparan Koin Tak Seimbang). *Science Tech: Jurnal Ilmu Pengetahuan Dan Teknologi*, 2(2), 61–69. <https://doi.org/10.30738/jst.v2i2.380>
- James, G., Daniela Witten, Hastie, T., & Tibshirani, R. (2013). *Springer Texts in Statistics An Introduction to Statistical Learning with application in R*.
- Journal, C., Mubarog, I., Setyanto, A., Sismoro, H., & Yoyakarta, U. A. (2019). *Sistem Klasifikasi pada Penyakit Breast Cancer dengan Menggunakan Metode Naïve Bayes*. 6(2), 109–118.
- Jumiati, E., Kamal, M. R., & Bootstrapping, S. (2015). *Integrasi sample bootstrapping pada k-nearest neighbor untuk klasifikasi herregistrasi calon mahasiswa baru. 1*.
- Junaedi, H., Budianto, H., Maryati, I., & Melani, Y. (2011). Data Transformation pada Data Mining. *Prosiding Konferensi Nasional Inovasi Dalam Desain Dan Teknologi-IDEaTech*, 7, 93–99. https://ideatech.stts.edu/proceeding/2011/12-000113_INF_Hartarto_p93-99.pdf
- Logo, J. F. B., Wantoro, A., & Susanto, E. R. (2020). Model Berbasis Fuzzy Dengan Fis Tsukamoto Untuk Penentuan Besaran Gaji Karyawan Pada Perusahaan Swasta. *Jurnal Teknoinfo*, 14(2), 124. <https://doi.org/10.33365/jti.v14i2.456>
- Mastutik, G., Alia, R., Rahniayu, A., Kurniasari, N., Rahaju, A. S., & Mustokoweni, S. (2015). Skrining Kanker Serviks dengan Pemeriksaan Pap Smear di Puskesmas Tanah Kali Kedinding Surabaya dan Rumah Sakit Mawadah Mojokerto. *Majalah Obstetri & Ginekologi*, 23(2), 54. <https://doi.org/10.20473/mog.v23i2.2090>
- Naufaldi, M. D., Gunawan, R., & Halim, R. (2020). *Pada Pasien Rawat Inap Di Rsud Raden Mattaher Jambi Tahun 2018-2020 Data dari Information Centre HPV on Berdasarkan Data dan Informasi Profil*. 48–58.
- Nugraha, P. A. (2013). *Perbandingan Metode Probabilistik Naive Bayesian Classifier dan Jaringan Syaraf Tiruan Learning Vector Quantization dalam Kasus Klasifikasi Penyakit Kandungan*. 2(2), 20–33.
- Rahman, M. F., Alamsah, D., Darmawidjadja, M. I., & Nurma, I. (2017). Klasifikasi Untuk Diagnosa Diabetes Menggunakan Metode Bayesian Regularization Neural Network (RBNN). *Jurnal Informatika*, 11(1), 36. <https://doi.org/10.26555/jifo.v11i1.a5452>
- Retnoningsih, E., & Pramudita, R. (2020). Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python. *Bina Insani Ict Journal*, 7(2), 156. <https://doi.org/10.51211/biict.v7i2.1422>
- Riana, D., Widyantoro, D. H., Latifah, T., & Mengko, R. (2013). Ekstraksi dan

- Klasifikasi Tekstur Citra Sel Nukleus Pap Smear. *Jurnal TICOM*,1(3), 62–7.
- Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. *IJCIT (Indonesian Journal on Computer and Information Technology)*, 5(1), 75–82. <https://doi.org/10.31294/ijcit.v5i1.7951>
- Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier dengan Metode Decision Tree Untuk Menganalisa Kelancaran Pembiayaan. *Jurnal TIM Darmajaya*, 02(01), 52–62.
- Sabransyah, M., Nasution, Y. N., & Tisna, D. (2017). *Aplikasi Metode Naive Bayes dalam Prediksi Risiko Penyakit Jantung Naive Bayes Method for a Heart Risk Disease Prediction Application*. 8, 111–118.
- Sari, E. R., & Alisah, E. (2012). Studi Tentang Persamaan Fuzzy. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, 2(2), 55–65. <https://doi.org/10.18860/ca.v2i2.2228>
- Sartika, D., Sensuse, D. I., Indo, U., Mandiri, G., & Komputer, F. I. (2017). *Perbandingan Algoritma Klasifikasi Naive Bayes , Nearest Neighbour , dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian*. 1(2), 151–161.
- Septiani, W. D., Studi, P., & Informatika, M. (2017). *Dan Naive Bayes Untuk Prediksi Penyakit Hepatitis*. 13(1), 76–84.
- Shanthi, P. B., Hareesha, K. S., & Kudva, R. (2022). *Automated Detection and Classification of Cervical Cancer Using Pap Smear Microscopic Images : A Comprehensive Review and Future Perspectives*. 20–41.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Sun, G., Li, S., Cao, Y., & Lang, F. (2017). *Cervical Cancer Diagnosis based on Random Forest*. 13(4), 446–457. <https://doi.org/10.23940/ijpe.17.04.p12.446457>
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). *Global Cancer Statistics 2020 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. 71(3), 209–249. <https://doi.org/10.3322/caac.21660>
- Sungkono, J. (2013). *Bootstrap pada r*. 84, 47–54.
- Suwandy, E., Bindan, H., Pranoto, E., & Dharma, A. (2020). *Analisa Metode Random Forest Tree dan K-Nearest Neighbor dalam Mendeteksi Kanker Serviks*. 3(2), 97–101.

- Taufiq, G. (2016). Implementasi Logika Fuzzy Tahani Untuk Model Sistem Pendukung Keputusan Evaluasi Kinerja Karyawan. *Jurnal Pilar Nusa Mandiri*, 12(1), 12–20. <http://pilar.nusamandiri.ac.id/index.php/pilar/article/view/58>
- Tütüncü, G. Y., & Kayaalp, N. (2015). Journal of Computational and Applied An Aggregated Fuzzy Naive Bayes Data Classifier. *Journal of Computational and Applied Mathematics*, 286, 17–27. <https://doi.org/10.1016/j.cam.2015.02.004>
- Ummah, I., Yannuansa, N., & Mufarrihah, I. (2021). Pengaruh Penentuan Domain, Fungsi Keanggotaan Dan Rule Dalam Membangun Sistem Fuzzy. *Jurnal Tecnoscienza*, 6(1), 165–175. <https://doi.org/10.51158/tecnoscienza.v6i1.607>
- Wahyuni, T., & Syswianti, D. (2018). *Sosialisasi Deteksi Dini Kanker Payudara Dan Kanker Serviks Melalui Webinar Berjudul “ Yuk Kenal Lebih Jauh Dengan Kanker Pembunuh Wanita Di Dunia ”*. 197–205.
- Wardani, A. R., Nasution, Y. N., & Amijaya, F. D. T. (2017). Aplikasi Logika Fuzzy Dalam Mengoptimalkan Produksi Minyak Kelapa Sawit Di PT. Waru Kaltim Plantation Menggunakan Metode Mamdani. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 12(2), 94. <https://doi.org/10.30872/jim.v12i2.651>
- Wibowo, S. (2015). Penerapan Logika Fuzzy Dalam Penjadwalan Waktu Kuliah. *Jurnal Informatika UPGRIS*, 1, 59–77.
- Wijaya, R. S. D., Adiwijaya, Andriyan B Suksmono, & Tati LR Mengko. (2021). Segmentasi Citra Kanker Serviks Menggunakan Markov Random Field dan Algoritma K-Means. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 139–147. <https://doi.org/10.29207/resti.v5i1.2816>
- Wirawan, I. N. T., & Eksistyanto, I. (2015). Penerapan Naive Bayes Pada Intrusion Detection System Dengan Diskritisasi Variabel. *JUTI: Jurnal Ilmiah Teknologi Informasi*, 13(2), 182. <https://doi.org/10.12962/j24068535.v13i2.a487>
- Yuliati, I. F., & Sihombing, P. R. (2021). *Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia Implementation of Machine Learning Method in Risk Classification on Low Birth weight in Indonesia*. 20(2), 417–426. <https://doi.org/10.30812/matrik.v20i2.1174>