

Optimasi *Decision Tree* Menggunakan *Particle Swarm Optimization* Untuk Klasifikasi Sel *Pap Smear*

Toni Arifin*¹, Rizal Rachman²

¹ Universitas ARS; Jl. Sekolah Internasional No 1-2, (022) 7100124 Universitas ARS.
Program Studi Teknik Informatika Universitas ARS Bandung.

² STMIK Nusa Mandiri, Jl. Keramat Raya No. 18, (021) 3906287 STMIK Nusa Mandiri.
Program Studi Sistem Informasi STMIK Nusa Mandiri Jakarta.
e-mail: *¹toni.arifin@ars.ac.id, ²rizal.rzc@nusamandiri.ac.id

Abstrak

Pap Smear adalah pemeriksaan keadaan sel-sel pada serviks di bawah mikroskop, kegunaan dari pemeriksaan *Pap Smear* adalah untuk menilai kesehatan organ kewanitaan dan juga sebagai cara mendeteksi kanker serviks sejak dini. Penelitian ini menyajikan klasifikasi sel *Pap Smear*. Data yang digunakan adalah dataset Herlev yang berjumlah 917 data. Penelitian ini bertujuan untuk mengetahui apakah penggunaan algoritma optimasi *particle swarm optimization* dapat meningkatkan kinerja dari algoritma *Decision tree* dalam mengklasifikasikan data Sel *Pap Smear*. Tahapan dari penelitian ini adalah *preprocessing*, *feature optimization*, *knowledge rules*, *evaluation* dan *performance report*. Hasil dari penelitian ini menunjukkan bahwa algoritma *Decision tree* menghasilkan akurasi sebesar 91.39 % dengan AUC 0.858, sedangkan penerapan algoritma *particle swarm optimization* pada *Decision tree* menghasilkan akurasi yang lebih baik yaitu sebesar 96.76 % dengan AUC 0.926..

Kata kunci— Herlev, *Pap Smear*, Klasifikasi, *Decision tree*, *particle swarm optimization*

Abstract

Pap Smear is an examination of the state of cells in the cervix under a microscope, the use of the *Pap smear* is to assess the health of female organs and also as a way to detect cervical cancer early. This research presents the classification of *Pap Smear* cells. The data used is the Herlev dataset which amounts to 917. This research aims to determine whether the use of *particle swarm optimization* algorithm can improve the performance of the *Decision tree* algorithm in classifying *Pap Smear Cell* data. The stages of this research are *preprocessing*, *feature optimization*, *knowledge rules*, *evaluation* and *performance report*. The results of this research are the *Decision Tree* algorithm produces an accuracy of 91.39% with AUC 0.858, While the application of *particle swarm optimization* algorithm in *Decision tree* resulted in better accuracy of 96.76% with AUC 0.926.

Keywords— Herlev, *Pap Smear*, Classification, *Decision tree*, *particle swarm optimization*

1. PENDAHULUAN

Kanker serviks adalah bentuk keganasan yang terjadi pada leher Rahim atau serviks [1], Kanker serviks yang ada pada bagian leher rahim merupakan penyebab kematian kedua didunia setelah kanker payudara [2] menurut data dari WHO kanker serviks adalah kanker yang paling umum diderita wanita [3]. Pada tahun 2018 diperkirakan ada 570.000 wanita di dunia didiagnosis kanker serviks dan sekitar 311.000 wanita meninggal karena kanker serviks [4], sedangkan menurut data dari *Global Cancer Observatory* pada tahun 2018 di Indonesia terdapat

32.426 kasus kanker serviks setiap tahunnya [5,6]. Di Indonesia 80% dari penderita kanker serviks memeriksakan diri dalam kondisi stadium lanjut, dan kurang lebih 94% pasien tersebut meninggal dalam kurun waktu 2 tahun [5]. Semua kasus kanker serviks disebabkan oleh *human papillomavirus* (HPV) [2,5,7]. *Human papillomavirus* (HPV) dikelompokkan menjadi 2 yaitu HPV yang beresiko rendah dan HPV yang beresiko tinggi, untuk tipe HPV yang beresiko rendah yaitu HPV 6, 11, 40, 42, 43, 44, 54 sedangkan HPV yang beresiko tinggi yaitu 16, 18, 26, 31, 33, 39, 45, 51, 52, 58, 82 [8], virus HPV tipe 16 dan 18 adalah tipe HPV yang berpotensi memunculkan kanker serviks [9].

Salah satu cara untuk mencegah dan menekan tingginya laju penderita kanker serviks adalah dengan melakukan pencegahan sejak dini, dengan melakukan pemeriksaan [5,7], pemeriksaan kanker serviks bisa dilakukan dengan cara melakukan pemeriksaan Pap Smear [10]. *Pap Smear* merupakan tes untuk mendeteksi kanker serviks pada wanita yang bertujuan untuk mengetahui ada atau tidaknya perubahan sel-sel rahim yang beresiko menjadi kanker [11]. Pengamatan sel *Pap Smear* di bawah mikroskop bukan tanpa kelemahan, kesalahan dalam pengamatan sel bisa terjadi [12,13], jika dikaitkan dengan perkembangan ilmu pengetahuan dan teknologi, teknik *machine learning* bisa jadi solusi untuk membantu ahli dalam identifikasi jenis sel *Pap Smear* [14,15].

Teknik *machine learning* atau nama lainnya adalah data mining [16], merupakan teknik untuk mengekstrak himpunan data dan menghasilkan sebuah pola agar lebih mudah untuk dianalisis dan diimplementasikan dalam sistem [16,17], Teknik *machine learning* dapat digunakan untuk klasifikasi jenis data [18]. Klasifikasi pada teknik *machine learning* bertujuan untuk memprediksi label atau *class* pada himpunan data [19]

Salah satu algoritma klasifikasi yang populer karena banyak digunakan dalam klasifikasi data dan banyak mempunyai kelebihan adalah algoritma *Decision tree* [17], algoritma ini dapat mengolah data numerik dan diskret [14], pada penelitian sebelumnya algoritma *Decision tree* lebih baik dari algoritma *Naïve bayes* [20], juga lebih baik dari algoritma K-NN [21] dan algoritma ini juga mempunyai kelemahan yaitu termasuk dalam kelompok algoritma klasifikasi tunggal, klasifikasi tunggal ini rentan dengan ketidakseimbangan data [16], klasifikasi tunggal lebih memprioritaskan *class mayoritas* dan mengabaikan informasi pada *class minoritas* [19]. Untuk Memaksimalkan algoritma *Decision tree* maka diperlukan teknik optimasi, teknik optimasi yang handal dalam optimasi fitur atau atribut yaitu algoritma optimasi *particle swarm optimization* [22]. *Particle swarm optimization* merupakan algoritma pencarian berbasis populasi, dan sangat efisien dalam memecahkan masalah optimasi [23] misalnya dalam identifikasi penyakit mata [22], sel *Pap Smear* [24], diabetes *retinopathy* [23], penilaian kinerja dosen [25], dan penyakit ibu hamil [26].

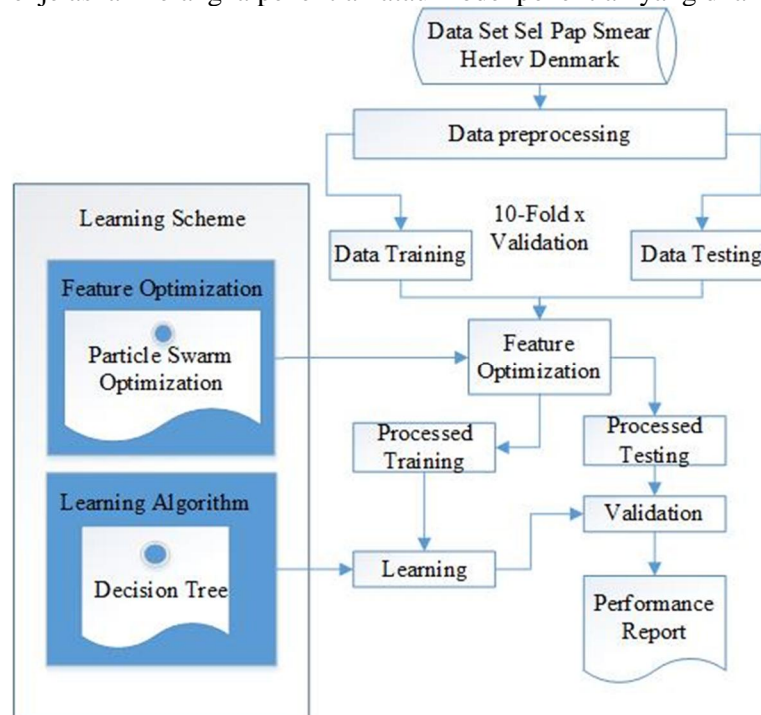
Beberapa penelitian sebelumnya yang mengangkat topik penelitian sel *Pap Smear* adalah penelitian Arifin, Riana dan Hapsari pada tahun 2014 [27] menggunakan algoritma GLCM untuk analisa tekstur dan *Decision tree* untuk klasifikasi. Penelitian ini menghasilkan nilai akurasi sistem sebesar 73% untuk klasifikasi sel *Pap Smear* menggunakan dataset Herlev Denmark. Penelitian lain yaitu yang dilakukan oleh Hidayatullah, Herliana, dan Arifin pada tahun 2016 [24] menggunakan dataset yang sama yaitu data sel Herlev Denmark dan menggunakan algoritma *Particle swarm optimization* dan *Naïve bayes* menghasilkan akurasi sistem sebesar 95,7%. Penelitian lainnya yaitu yang dilakukan oleh Dewi dan Sariasih pada tahun 2019 [28] menggunakan dataset Herlev Denmark dengan algoritma *Sample bootstrapping* dan *Naïve bayes* dengan akurasi sistem yaitu mencapai 93,24 % untuk 2 kelas yaitu normal dan abnormal.

Berdasarkan dari penelitian sebelumnya dengan menggunakan dataset Herlev Denmark untuk klasifikasi data sel *Pap Smear* hasil akhir masih bisa ditingkatkan menjadi lebih baik. Penelitian ini bertujuan untuk melakukan penerapan algoritma *Decision tree* yang dioptimasi dengan menggunakan algoritma *Particle swarm optimization* untuk meningkatkan akurasi

klasifikasi sel *Pap Smear* normal dan abnormal. Hasil dari penelitian ini diharapkan dapat membantu ahli patologi dalam mengidentifikasi jenis sel nukleus normal dan abnormal.

2. ALGORITMA PENELITIAN

Dari pemaparan pendahuluan dan *state-of-the-art* sebelumnya maka dibutuhkan model penelitian untuk menggambarkan *rule* penelitian dari proses awal hingga proses akhir. Di bawah ini Gambar 1 menjelaskan kerangka penelitian atau model penelitian yang dilakukan.



Gambar 1. Kerangka Penelitian

2.1 Dataset

Penelitian ini menggunakan data dari penelitian sebelumnya yaitu dari Jantzen, Norup, Dounias dan Bjerregaard pada tahun 2005 [29], jumlah data terdiri dari 917 data citra yang sudah di ekstrak dan menghasilkan basis data atau dataset berupa excel yang terdiri dari 20 atribut. Diantaranya adalah *Kerne_A*, *Cyto_A*, *K/C*, *Kerne_Ycol*, *Cyto_Ycol*, *KerneShort*, *KerneLong*, *KerneElong*, *KerneRund*, *CytoShort*, *CytoLong*, *CytoElong*, *CytoRund*, *KernePeri*, *CytoPeri*, *KernePos*, *KerneMax*, *KerneMin*, *CytoMax*, *CytoMin*.

2.2 Data Preprocessing

Pada tahap ini data dianalisa dan di olah dengan menggunakan teknik data *cleansing* [16], lalu data yang sudah diolah dibagi ke dalam 2 bagian yaitu data *training* dan data *testing* dengan menggunakan teknik *10 fold cross validation* [16,18,30].

2.3 Particel Swarm Optimization

Gambar 1 kerangka penelitian *Particel Swarm Optimization* terdapat pada bagian *learning Scheme*, tahapan ini merupakan tahapan optimasi data atau atribut untuk data yang digunakan [31]. *Algoritma Particel Swarm Optimization* [31], [18,30].

1. Asumsi bahwa ukuran kelompok atau kawan (jumlah partikel) N .
2. Bangkitkan populasi awal X dengan rentang $X(B)$ dan $X(A)$ secara random sehingga didapat $X1, X2, \dots, XN$. Setelah itu, untuk mudahnya, partikel j dan kecepatannya pada iterasi i dinotasikan sebagai $X(i)j$ dan $V(i)j$ Sehingga partikel-partikel awal ini akan menjadi $X1(0), X2(0), \dots, XN(0)$.

3. Hitung kecepatan dari semua partikel Set iterasi $i=1$.
4. Pada iterasi ke- i , temukan 2 parameter penting untuk setiap partikel j .
5. Hitung kecepatan partikel j pada iterasi ke i

$$v_{i,m} = w.v_{i,m} + c1 * R * (pbest_{i,m} - x_{i,m}) + c2 * R * (gbest_m - x_{i,m}) \quad (1)$$

6. Hitung posisi atau koordinat partikel j pada iterasi ke- i dengan cara

$$x_{id} = x_{i,m} + v_{i,m} \quad (2)$$

2.4 Decision Tree

Decision Tree pada Gambar 1 kerangka penelitian adalah tahapan pencarian pola klasifikasi dari dataset yang digunakan [16,17].

1. Menentukan atribut, diambil dari dataset sel Pap Smear.
2. Menentukan atribut yang terpilih lebih dahulu dengan menghitung nilai entropi

$$entropi(y) = \sum_{i=1} - p_i * \log_2 p_i \quad (3)$$

3. Menghitung nilai information gain terbesar
4. Menghitung information gain dari output y data yang dikelompokkan berdasarkan atribut A :
 Hitung Gain (y, A)

$$Gain(y, A) = Entropi(y) - \frac{y_c}{y} Entropi(y_c) \quad (4)$$

5. Ulangi langkah ke-2 hingga semua atribut terpecah

2.5 Validasi

Tahap Validasi pada Gambar 1 bertujuan untuk mengetahui tingkat akurasi dari pola yang dihasilkan [16]. Teknik validasi yang digunakan pada penelitian ini adalah *Confusion matrix*, ini adalah persamaan dari validasi menggunakan *Confusion matrix*.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Keterangan:

TP = Nilai prediksi benar dan nilai sebenarnya benar.

FN = Nilai prediksi salah dan nilai sebenarnya benar.

FP = Nilai prediksi benar dan nilai sebenarnya salah.

TN = Nilai prediksi salah dan nilai sebenarnya salah.

3. HASIL DAN PEMBAHASAN

Hasil dari penelitian yang sudah dilakukan, diantaranya Tabel 2 menjelaskan mengenai hasil dari algoritma *Particel Swarm Optimization* yang berupa bobot dari 20 atribut yang digunakan, hanya ada 10 atribut yang menghasilkan nilai lebih dari 0.5 yaitu Kerne A, Cyto A, KerneElong, KerneRund, CytoLong, CytoRund, KenerPeri, KerneMax, CytoMax dan CytoMin, atribut ini yang akan digunakan dalam tahapan klasifikasi untuk menghasilkan pola atau *rule*, sedangkan pada Gambar 2 adalah hasil dari Algoritma *Decision tree* yang berupa *rule* klasifikasi dimana atributnya disesuaikan dengan hasil dari algoritma *Particel Swarm Optimization*, dan Tabel 3 merupakan hasil dari validasi dengan menggunakan teknik *Confusion*

Matrix untuk kategori kelas normal dan abnormal, pada tabel juga dijelaskan perbandingan hasil dari algoritma yaitu untuk model yang hanya menggunakan algoritma *Decision tree* menghasilkan nilai akurasi sebesar 91,39 % dan AUC 0.858 dan hasil dari *Decision tree* dengan *Particel Swarm Optimization* menghasilkan akurasi sebesar 96,76% dan AUC 0.926.

Tabel 2. Hasil dari Algoritma *Particel Swarm Optimization*

No	Atribut	Weight	No	Atribut	Weight
1	Kerne_A	1	11	CytoLong	1
2	Cyto_A	0.955	12	CytoElong	0
3	K/C	0.083	13	CytoRund	0.982
4	Kerne_Ycol	0.373	14	KernePeri	1
5	Cyto_Ycol	0	15	CytoPeri	0
6	KerneShort	0	16	KernePos	0.005
7	KerneLong	0	17	KerneMax	1
8	KerneElong	1	18	KerneMin	0
9	KerneRund	1	19	CytoMax	0.819
10	CytoShort	0	20	CytoMin	1

```

Kerne_A > 1464
| Cyto_A > 63854.688: normal {normal=4, abnormal=0}
| Cyto_A ≤ 63854.688
| | Cyto_A > 36905.688
| | | Kerne_A > 3711: abnormal {normal=0, abnormal=8}
| | | Kerne_A ≤ 3711: normal {normal=14, abnormal=0}
| | Cyto_A ≤ 36905.688
| | | KerneElong > 1.129: normal {normal=1, abnormal=1}
| | | KerneElong ≤ 1.129
| | | | Kerne_A > 1628.375
| | | | | KernePeri > 148.938
| | | | | | CytoMin > 35.500
| | | | | | | Kerne_A > 3352.312: abnormal {normal=1, abnormal=305}
| | | | | | | Kerne_A ≤ 3352.312
| | | | | | | | Cyto_A > 26644.812: normal {normal=2, abnormal=0}
| | | | | | | | Cyto_A ≤ 26644.812: abnormal {normal=28, abnormal=317}
| | | | | | | | CytoMin ≤ 35.500: abnormal {normal=1, abnormal=3}
| | | | | | | KernePeri ≤ 148.938: abnormal {normal=1, abnormal=2}
| | | | | Kerne_A ≤ 1628.375
| | | | | | KerneMax > 79.500: normal {normal=3, abnormal=0}
| | | | | | KerneMax ≤ 79.500
| | | | | | | Kerne_A > 1535.125
| | | | | | | | KerneRund > 0.476: normal {normal=5, abnormal=2}
| | | | | | | | KerneRund ≤ 0.476: abnormal {normal=0, abnormal=5}
| | | | | | | | Kerne_A ≤ 1535.125: abnormal {normal=0, abnormal=9}
Kerne_A ≤ 1464
| CytoLong > 57.595: normal {normal=182, abnormal=21}
| CytoLong ≤ 57.595: abnormal {normal=0, abnormal=2}

```

Gambar 2. Hasil dari Algoritma *Decision Tree*

Tabel 3. *Confusion Matrix* Kelas Normal dan Abnormal

No	Algoritma	Hasil	AUC
1	Decision Tree	91.39%	0.858
2	Decision Tree + Particel Swarm Optimization	96,76%	0.926

Hasil pada Tabel 3 menunjukkan bahwa algoritma *Particel Swarm Optimization* dapat mengatasi masalah dari ketidakseimbangan data yang dihasilkan oleh *Decision tree*, dan efisien dalam optimasi data seperti pada penelitian identifikasi penyakit mata [22], sel Pap Smear [24], diabetes retinopathy [23], penilaian kinerja dosen [25], dan penyakit ibu hamil [26], penelitian ini juga menghasilkan akurasi yang lebih baik dari [24,27,28] dengan menghasilkan akurasi sebesar 96,76 % dan termasuk dalam kategori *excellent classification* karena menghasilkan AUC 0.926 untuk identifikasi sel Pap Smear normal dan abnormal.

4. KESIMPULAN

Berdasarkan dari eksperimen yang sudah dilakukan pada Tabel 3, algoritma *Decision tree* menghasilkan nilai akurasi sebesar 91.39 % dan hasil dari optimasi algoritma *Decision tree* menggunakan *Particel Swarm Optimization* dapat menghasilkan akurasi yang cukup baik yaitu sebesar 96,76 %, hal ini menunjukkan bahwa penerapan algoritma *Particel Swarm Optimization* dan *Decision tree* lebih baik dari penelitian Hidayatullah, Herliana, dan Arifin pada tahun 2016 [24], Arifin, Riana dan Hapsari pada tahun 2014 [27], dan Dewi dan Sariasih pada tahun 2019 [28].

UCAPAN TERIMA KASIH

Ucapan terima kasih juga kami ucapkan kepada Kementerian Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia untuk pendanaan penelitian ini.

DAFTAR PUSTAKA

- [1] P. R. Evriarti and A. Yasmon. 2019, “*Patogenesis Human Papillomavirus (HPV) pada Kanker Serviks*,” *Litbang Kemkes*, Vol. 8, No. 1.
- [2] M. Fauza, 2019, “*Faktor yang Berhubungan Dengan Deteksi Dini Kanker Serviks Metode IVA di Puskesmas Kota Padang*,” Vol. 14, No. 1.
- [3] WHO, 2020, “*Cancer*,” [Online]. Available: https://www.who.int/health-topics/cancer#tab=tab_1. [Accessed: 15-Jan-2020].
- [4] WHO, 2020, “*Cervical Cancer*,” [Online]. Available: https://www.who.int/health-topics/cervical-cancer#tab=tab_1. [Accessed: 10-Mar-2020].
- [5] A. Wardani, 2019, “*Kasus Kanker Serviks Indonesia Terbanyak Kedua di Dunia*,”
- [6] G. C. Observatory, 2019. “*Cervical Cancer*,” [Online]. Available: <https://gco.iarc.fr/>. [Accessed: 01-Jun-2020].

-
- [7] K. Adrian, 2020, “Penyebab Kanker Serviks dan Faktor Risiko yang Perlu Diketahui.”
- [8] D. Murtiastutik, G. Mastutik, A. Rahniayu, A. Arista, and T. Setyaningrum, 2019, “The Genotype of Human Papilloma Virus of Male Patients With Anogenital,” Vol. 55, No. 2, pp. 100–106.
- [9] A. Djie, 2019, “Virus HPV, Salah Satu Penyebab Kanker Serviks,”
- [10] A. N. Sholihah and E. Sulistyorini, 2015, “Hubungan Antara Sikap Pencegahan Kanker Serviks Dengan Minat Deteksi Dini Menggunakan Inspeksi Visual Asam Asetat pada Wanita Usia Subur Di Rw Iv Desa Cangkol Mojolaban Sukoharjo Tahun 2015,” pp. 102–116.
- [11] L. Tan, 2020, “Pap Smear,” [Online]. Available: [https://www.sehatq.com/tindakan-medis/pap-smear-siapa-saja-yang-membutuhkannya#:~:text=Pap Smear adalah Sebuah tes,serviks\) yang Berisiko Menjadi Kanker.](https://www.sehatq.com/tindakan-medis/pap-smear-siapa-saja-yang-membutuhkannya#:~:text=Pap%20Smear%20adalah%20Sebuah%20tes,serviks) yang Berisiko Menjadi Kanker.) [Accessed: 06-Jun-2020].
- [12] I. Muhimmah, 2013, “Analysis of Features to Distinguish Epithelial Cells and Inflammatory Cells in Pap Smear Images,” No. Bmei, pp. 519–523.
- [13] I. Muhimmah, R. Kurniawan, and Indrayanti, 2012, “Automated Cervical Cell Nuclei Segmentation Using Morphological Operation and Watershed Transformation,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, pp. 163–168.
- [14] T. Arifin, 2017, “Implementasi Algoritma PSO Dan Teknik Bagging Untuk Klasifikasi Sel Pap Smear,” *J. Inform.*, Vol. 4, No. 2, pp. 155–162.
- [15] T. Arifin, 2015, “Implementasi Metode K-Nearest Neighbor Untuk Klasifikasi Citra Sel Pap Smear Menggunakan Analisis Tekstur Nukleus,” *J. Inform.*, Vol. 2, No. 1, pp. 287–295.
- [16] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, 2017, *Practical Machine Learning Fourth Edition*. Elsevier MK.
- [17] V. Kotu and B. Deshpande, 2015, *Predictive Analytics and Data Mining*. Elsevier MK.
- [18] D. L. Olson and D. Wu, 2020, *Predictive Data Mining Models*, Second Edi. Springer.
- [19] N. Ye, 2014, *Data Mining Theories, Algorithms, and Examples*, CRC Press.
- [20] T. Rosandy, 2016, “Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4 . 5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus: KSPPS/ BMT AL-FADHILA),” Vol. 02, No. 01, pp. 52–62.
- [21] S. Wahyuningsih, D. R. Utari, U. B. Luhur, D. Tree, and K. Validation, 2018, “Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decision Tree Untuk Prediksi Kelayakan Pemberian Kredit,” pp. 8–9.
- [22] T. Arifin and A. Herliana, 2020, “Optimasi Decision Tree Menggunakan Particle Swarm Optimization Untuk Identifikasi Penyakit Mata Berdasarkan Analisis Tekstur Optimizing

- Decision Tree Using Particle Swarm Optimization to Identify Eye Diseases,*” Vol. 8, No. January, pp. 59–63.
- [23] T. Arifin and A. Herliana, 2018, “*Optimasi Metode Klasifikasi Dengan Menggunakan Particle Swarm Optimization Untuk Identifikasi Penyakit Diabetes Retinopathy,*” *khazanah Inform.*, Vol. 4, No. 2, pp. 77–81.
- [24] T. Hidayatulloh, A. Herliana, and T. Arifin, 2016, “*Klasifikasi Sel Tunggal Pap Smear Berdasarkan Analisis Fitur Berbasis Naïve Bayes Classifier,*” Vol. IV, No. 2, pp. 186–193.
- [25] M. Sihombing and E. Manik, 2019, “*Jurnal Teknologi, Kesehatan dan Ilmu Sosial Penerapan Fuzzy Dengan Particle Swarm Optimization (PSO) Jurnal Teknologi, Kesehatan dan Ilmu Sosial,*” Vol. 1, No. 1, pp. 1–4.
- [26] A. Byna, “*Penerapan Optimasi Pso Untuk Meningkatkan Akurasi Algoritma Id3 pada Prediksi Penyakit Ibu Hamil,*” pp. 65–70.
- [27] T. Arifin, D. Riana, and G. I. Hapsari, 2014, “*Klasifikasi Statistik Tekstur Sel Pap Smear Dengan Decesion Tree,*” *J. Inform.*, Vol. 1, No. 1.
- [28] Y. N. Dewi and F. A. Sariasih, 2019, “*Metode Sample Bootstrapping Untuk Meningkatkan Performa,*” Vol. 12, No. 1.
- [29] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, 2015, “*Pap-smear Benchmark Data For Pattern Classification Pap-smear Benchmark Data For Pattern Classification,*” No. March.
- [30] M. North, 2012, *Data Mining for the Masses*. A Global Text Project Book.
- [31] Y. Zhang, S. Wang, P. Phillips, and G. Ji, 2014, “*Binary PSO With Mutation Operator For Feature Selection Using Decision Tree Applied To Spam Detection,*” *Knowledge-Based Syst.*, Vol. 64, pp. 22–31.