

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

# Klasifikasi Gejala Penyakit Coronavirus Disease 19 (COVID-19) Menggunakan Machine Learning

Suci Anggraini<sup>1</sup>, Muhamad Akbar\*<sup>2</sup>, Alex Wijaya<sup>3</sup>, Hadi Syaputra<sup>4</sup>, Muhammad Sobri<sup>5</sup>

1,2,3,4,5 Informatics Department , Bina Darma University, Palembang, Indonesia Email: rainisuci@gmail.com¹, muhamad.akbar@binadarma.ac.id², alex\_wj@binadarma.ac.id³, hadisyaputra@binadarma.ac.id⁴, sobri@binadarma.ac.id⁵

#### Abstrak

Coronavirus Disease 19 (COVID-19) merupakan virus baru yang menyebabkan infeksi saluran pernapasan. Virus ini berasal dari hewan yang dapat menular pada manusia dengan percikan air liur. Menurut data epidemiologi rata-rata pasien terjangkit virus ini berusia 15-80 tahun. Virus ini memiliki masa inkubasi 2-14 hari yang mempunyai gejala awal yaitu deman tinggi, sesak nafas, batuk pilek. Indonesia memiliki 2 kasus pertama pada 2 maret 2020. Permasalahan yang diangkat dalam penelitian ini adalah bagaimana mengklasifikasi resiko terjangkit virus covid-19 dari gejala yang ditimbulkan. Tujuan penelitian ini untuk mengetahui tingkat resiko terjangkit virus covid-19 berdasarkan instrumen yang digunakan dari metode Knowledge Discovery in Database yang terdiri dari 5 tahapan yaitu selection, pre-processing, transformation, data mining, dan evaluation. Dataset yang digunakan peneliti diambil dari web resmi kaggle.com. Penelitian ini menggunakan 4 (empat) algoritma yaitu K-Nearest Neighbor (K-NN), Neural Network (NN), Random Forest (RF), dan Naive Bayes dengan bantuan tool rapidminer. Values dataset antara lain tingkat rendah 25,98%, tingkat sedang 54,33%, dan tingkat tinggi 19,69%. Nilai akurasi pada dataset dengan 127 data pasien terjangkit covid-19 menggunakan algoritma k-nearest neighbor memperoleh 57,89%, neural network memperoleh 73,68%, random forest memperoleh 68,42%, naive bayes memperoleh 65,38%. Pada penelitian ini algoritma klasifikasi Neural Network memberikan nilai akurasi yang tertinggi.

Kata Kunci: Covid-19, Klasifikasi, Rapidminer, K-Nearest Neighbor, Neural Network, Random Forest, Naive Bayes, Knowledge Discovery in Database, kaggle.com, Machine Learning.

#### 1. PENDAHULUAN

World Health Organization (WHO) memberitahukan kasus baru Pneumonia di kota Wuhan, Hubei, China yang mengidentifikasi jenis baru novel coronavirus. Nama Coronavirus Disease 2019 resmi ditetapkan oleh WHO [1]. Menurut ahli



Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

virologi dari China, virus covid-19 ini merupakan virus yang berbeda dengan Severe Acute Respiratory Syndrome Associated Coronavirus (SARS-COV2) yang muncul di Guandong, China tahun 2003 tetapi memiliki gajala yang sama [1]. Tingkat penyebaran Covid-19 lebih luas dibandingkan dengan SARS namun tingkat kematian SARS mencapai 9,6% dibanding tingkat kematian Covid-19 yang masih dibawah 5% [1]. Homologi Covid-19 mempunyai ciri-ciri DNA yang mirip hingga mencapai 85% dengan kelelawar SARS. Penularan virus ini dari hewan ke manusia disebut Transmisi Zoonosis dan dapat tertular dari manusia ke manusia dengan berkontak langsung atau terkena percikan liurnya [2]. Dari data pertama di wuhan 15% menunjukan kasus fatal usia diatas 80 tahun, 8,0% berusia 70 tahun, 1% anak dibawah 15 tahun. Sedangkan kasus ringan dan berat yang mempunyai penyakit bawaan 49,0%[1].

Ada 4 (empat) kategori penyebutan Covid-19 dalam pedoman WHO dan Penelitian Stigma Orang Positif Covid-19 yaitu Orang Dalam Pemantauan, Pasien Dalam Pengawasan, Orang Tanpa Gejala, dan Positif [3]. Pada tanggal 28 desember 2020 di indonesia sebanyak 719.219 ribu orang dinyatakan terkonfirmasi Covid-19 dan kasus meninggal mencapai 21 ribu jiwa [4]. Menurut penelitian Tidora tentang kelompok beresiko tinggi terinfeksi Covid-19 ialah yang mempunyai penyakit bawaan, seorang perokok, dan lansia [5]. Diindonesia sendiri telah melakukan upaya untuk memperlambat penyebaran Covid-19 dengan cara melakukan 3M, menjaga jarak, memakai masker, dan mencuci tangan [6]. Dalam penelitian ini menggunakan 4 (empat) algoritma yang paling sering digunakan yaitu K-Nearest Neighbor yang merupakan algoritma terbaik [7]. Random Forest merupakan algoritma yang telah teruji dapat menghasilkan nilai akurasi tinggi [8]. Neural Network dapat menghasilakan nilai akurasi hampir mencapai 100% [9]. Sedangkan Naive Bayes memberikan nilai akurasi diatas 75% [10]. Peneliti menggunakan software rapidminer untuk mempermudah dalam menganalisis [11].

## 2. METODOLOGI PENELITIAN

## 2.1 Knowledge Discovery in Database (KDD)

Pada tahap penelitian peneliti menggunakan metode KDD untuk mengklasifikasi pasien terjangkit covid-19 dengan mencari pola data pada dataset yang sudah diunduh di web resmi kaggle.com. Dibawah ini merupakan tahapan pelaksanaan dari metode KDD sebagai berikut:

#### 1. Selection

Pemilihan data yang relavan untuk dikelola sesuai dengan tujuan penelitian. Dari pemilihan data yang telah dilakukan akan disimpan terpisah dan dikembalikan pada database.

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

## https://journal-computing.org/index.php/journal-sea/index

# 2. Pre-processing

Berfungsi untuk pembersihan data seperti data duplikat permasalahan yang lain dan untuk memperbanyak data. relavan sebelum digunakan.

## 3. Transformation

Fitur presentase data yang terpilih berupa kode, dan bergantung pada pola informasi.

#### 4. Data Mining

Pencarian pola dengan metode tertentu, yang artinya metode ini bergantung pada alat dan algoritma yang digunakan. Peneliti menggunakan software RapidMiner dan empat Algoritma.

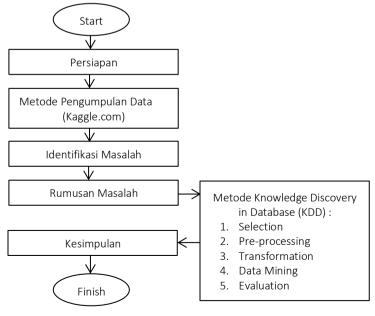
#### 5. Evaluation

Hasil dari proses pengujian vaitu pemeriksaan pola. Dilakukan untuk mengetahui apakah pola tersebut sama dengan apa yang diharapkan atau hasilnya bertentangan.

## 2.2 Metode Pengumpulan Data

Penelitian klasifikasi ini peneliti menggunakan data Gejala Peyakit Coronavirus Disease 19 yang diunduh pada situs kaggle.com Covid-19 Patient Symptons, 2020. url https://www.kaggle.com/bitsofishan/covid19-patient-symptoms Dataset tersebut berisi 127 rows, 21 columns, dan 21 atribut.

#### 2.3 Alur Penelitian



Gambar 1. Alur Pelaksanaan

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

Alur pelaksanaan pada penelitian merupakan hal yang penting, agar penelitian yang akan dilaksanakan dapat berjalan dengan baik dan sesuai dengan keinginan. Dan dapat juga mempermudah peneliti untuk memahami proses yang akan dilaksanakan. Pada penelitian ini akan melalui tahapan tahapan : Persiapan yang dilakukan adalah mengenali apa yang akan diteliti dengan mempelajari masukan jurnal terdahulu dan studi litelature lainnya yang dapat menjadi sumber tentang judul yang akan diteliti. Tahapan yang kedua peneliti melakukan pengambilan data dari situs resmi Kaggle.com yang merupakan web pengumpulan data untuk bahan pembelajaran, data yang diambil berasal dari india yang mempunyai 127 data pasien yang terjangkit Covid-19. Kemudian tahap ketiga mengidentifikasi masalah yaitu dengan memilih metode dan algoritma yang akan digunakan pada penelitian, peneliti menggunakan software rapidminer dengan empat algoritma klasifikasi yaitu, K-Nearest Neighbor, Naïve Bayes, Random Forest, dan Neural Network. Yang ke empat yaitu mencari solusi dari permasalahan yang telah ditemukan. Selanjutnya pengolahan data dengan menggunakan K-Nowledge Discovery in Database (KDD) yang meliputi selection pre-processing transformasi, data mining, dan evaluation. Untuk mengetahui kelompok tingkat rendah, sedang, tinggi pasien terjangkit covid-19. Kemudian memberikan hasil serta pembahasan melakukan pengujian menggunakan tools dan algoritma yang telah dipilih. Dan terakhir saran, pada tahap ini peneliti akan memberikan kesimpulan dari hasil uji data mining serta memberikan sedikit saran untuk penelitian selanjutnya.

## 3. HASIL DAN PEMBAHASAN

#### 3.1. Hasil Perhitungan

Dari penelitian yang telah dilakukan menggunakan rapidminer terdapat tiga label klasifikasi yaitu tingkat rendah (low risk), tingkat sedang (medium risk), tingkat tinggi (high risk) menggunakan empat algoritma k-nearest neighbor, random forest, neural network, dan naive bayes dengan data (127 pasien).

Random Forest dan Neural Network menggunakan data random dikarenakan algoritma ini memiliki pengaturan tersendiri dalam pengelolaan label pada aplikasi rapidminer sedangkan K-Nearest Neighbor dan Naive Bayes menggunakan data training dan testing dimulai dari 50:50 sampai 80:20. Hasil Akurasi dari keempat algoritma tersebut dapat dilihat pada table di bawah ini:

Table 1 Hasil Akurasi Algoritma Random Forest dan Neural Network

Algoritma	Nilai Akurasi
Random Forest	68,42%
Neural Network	73,68%

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

Table 2 Hasil Akurasi Algoritma K-Nearest Neighbor dan Naive Bayes

Training & Testing	Nilai Akurasi K-Nearest Neighbor	Nilai Akurasi Naive Bayes	
50 : 50	49,21%	58,73%	
60 : 40	52,94%	47,06%	
70 : 30	57,89%	63,16%	
80 : 20	42,31%	65,38%	

Algoritma K-Nearest Neighbor (K-NN) dan Naive Bayes diambil pada nilai akurasi tertinggi yaitu pada algoritma K-NN nilai akurasi tertinggi terdapat pada training dan testing 70:30 (57,89%). Untuk Algoritma Naive Bayes terdapat nilai akurasi tertinggi pada training dan testing 80:20 (65,38%). Dari keempat algoritma yang telah diuji diambil nilai akurasi tertinggi, dapat dilihat pada table 3 berikut ini :

Table 3 Nilai akurasi Tertinggi sampai Terendah

Algoritma	Nilai Akurasi
Neural Network	73,68%
Random Forest	68,42%
Naive Bayes	65,38%
K-Nearest Neighbor	57,89%

Pada klasifikasi gejala penyakit covid-19. Algoritma tertinggi yaitu Neural Network 73,68%. Algoritma yang mendapatkan nilai akurasi terendah yaitu K-Nearest Neighbor nilai akurasi sebesar 57,89%. Hal ini menunjukan bahwa data training dan testing pada dataset dapat mempengaruhi performa klasifikasi pada setiap Algoritma.

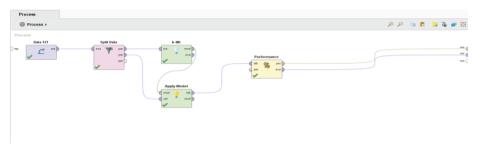
#### 3.2. K-Nearest Neighbor (K-NN)

K-NN merupakan salah satu algoritma klasifikasi terbaik (Wu et al. 2007) [7], yang cara kerjanya menentukan jarat paling dekat [12]. K-Nearest Neighbor mempunyai kelebihan yaitu mudah untuk dipelajari, algoritma yang sederhana, dan kinerja yang baik untuk data besar [13]. Pada penelitian ini data sebanyak 127 baris dengan 21 kolom, dan split data pada Machine Learning untuk membagi data latih (70%) dan uji (30%) setelah itu digunakan algoritma K-Nearest Neighbor sebagai metode klasifikasi, dan dihubungkan dengan split data algoritma yang digunakan pada Apply Model untuk mengetahui data yang tidak mempunyai label. Performance untuk mengevaluasi hasil kerja yang akan mendapatkan nilai akurasi sebesar 57,89%.

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

Berikut aplikasi yang sudah diterapkan dapat dilihat pada Gambar 2 dan Gambar 3.



Gambar 2 Model sub proses klasifikasi K-NN



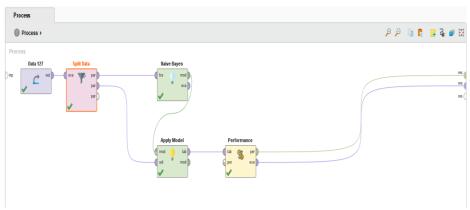
Gambar 3 Hasil akurasi K-NN

## 3.3. Naïve Bayes (NB)

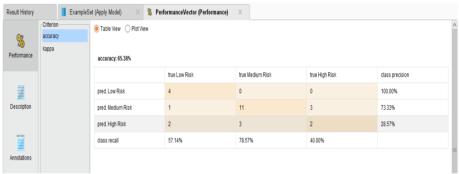
Naive Bayes terkenal sebagai algoritma yang mempunyai kinerja baik (Siddik et al.)[14] dan merupakan algoritma yang mampu memperkirakan variabel bersifat bebas [15]. Penelitian ini data sebanyak 127 baris dengan 21 kolom, dan split data pada Machine Learning untuk membagi antara data latih (80%) dan uji (20%) setelah itu digunakan algoritma Naive Bayes sebagai metode klasifikasi, dan dihubungkan dengan split data algoritma yang digunakan pada Apply Model untuk mengetahui data yang tidak mempunyai label. Performance untuk mengevaluasi hasil kerja yang akan mendapatkan nilai akurasi sebesar 65,38%. Berikut aplikasi yang sudah diterapkan dapat dilihat pada Gambar 4 dan Gambar 5.

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index



Gambar 4 Model Sub Proses Naive Bayes



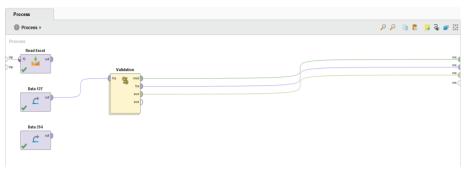
Gambar 5 Hasil Akurasi dari Naive Bayes

#### 3.4. Random Forest (RF)

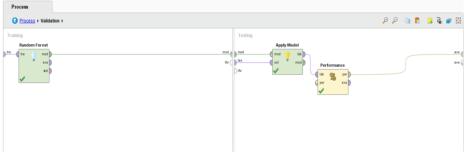
Pada penelitian yang telah di lakukan sebelumnya menggunakan algoritma Random forest menyatakan algoritma ini dapat menghasilkan nilai akurasi yang tinggi [8]. Algoritma ini merupakan keturunan dari ID3 di desain oleh J.Ross Quinlan [16]. Random Forest memiliki kelebihan antara lain, memiliki kinerja yang baik, dapat membentuk error yang kecil, dan memberikan training lebih besar [17]. Hasil Penelitian dari data sebanyak 127 baris dengan 21 kolom, dan split validation pada rapidminer untuk mengukur kinerja model dengan pembagian data secara random, setelah itu digunakan algoritma Random Forest sebagai metode klasifikasi, dan dihubungkan dengan Apply Model untuk mengetahui data yang tidak mempunyai label. Performance untuk mengevaluasi hasil kerja yang akan mendapatkan nilai akurasi sebesar 68,42%. Berikut aplikasi yang sudah diterapkan dapat dilihat pada Gambar 6, Gambar 7, dan Gambar 8.

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index



Gambar 6 Model Sub Proses Random Forest



Gambar 7 Model Sub Validation Random Forest



Gambar 8 Hasil Akurasi dari Random Forest

## 3.5. Neural Network (NN)

Algoritma Neural Network pada penelitian penyakit jantung memberikan nilai akurasi 91,45% [9] dan merupakan algoritma yang sering digunakan dalam latihan supervised [18]. Neural Network pada software rapidminer mempunyai formula yang dirangkum sebagai training learning [18] serta mempunyai keutamaan dalam prediksi dan kecepatan missing data [19]. Hasil Penelitian dari data sebanyak 127 baris dengan 21 kolom, dan split validation pada rapidminer untuk mengukur kinerja model dengan pembagian data secara random. Dan digunakan algoritma Neural Network sebagai metode

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

klasifikasi, dan dihubungkan dengan Apply Model untuk mengetahui data yang tidak mempunyai label. Performance untuk mengevaluasi hasil kerja yang mendapatkan nilai akurasi sebesar 73,68%. Berikut aplikasi yang sudah diterapkan dapat dilihat pada Gambar 9, Gambar 10, dan Gambar 11.



Gambar 9 Model Sub Proses Neural Network



Gambar 10 Model Sub Validation Neural Network



Gambar 11 Hasil Akurasi dari Neural Network

## 3.6. Pembahasan Gejala Covid-19

Dari dataset yang diunduh pada situs kaggle.com berupa 127 data pasien terjangkit covid-19 telah di klasifikasi menggunakan software rapidminer dan menunjukan hasil absolute count pada tingkat rendah sebesar 25,98%, tingkat sedang 54,33%, dan tingkat tinggi sebesar 19,69%. Dapat dilihat pada Gambar 12 berikut:

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index



Gambar 12 Hasil Nominal Values

Gejala Rendah: Telah dipilih beberapa gejala yang paling banyak muncul pada pasien positif Covid-19 yang telah di klasifikasi. Ada 5 gejala yaitu memiliki suhu tubuh ≥ 39 C, batuk kering, sakit tenggorokan, perkembangan gejala, perubahan pada nafsu makan, dapat dilihat pada table 4 Sekitar 0,27% pasien merasakan tiga gejala yang sama.

Table 4 Dari Data Klasifikasi Gejala Rendah

No Umur	Hmur	Suhu Tubuh	Batuk Kering	Sakit	Perkembangan	Perubahan
	Offici			Tenggorokan	Gejala	Nafsu Makan
1	58	1	1	1	0	0
2	44	1	1	0	1	1
3	37	1	1	1	0	1
4						
33	33	1	0	1	1	1

❖ Gejala Sedang: Dipilih beberapa gejala yang paling banyak muncul dari data pasien covid-19 yang telah di klasifikasi antara lain suhu tubuh ≥ 39 C, batuk kering, sakit tenggorokan, kantuk, memiliki riwayat perjalanan, dapat dilihat pada table 5 Sekitar 0,14% pasien merasakan empat gejala sekaligus.

Table 5 Dari Data Klasifikasi Geiala Sedang

No	Umur	Suhu Tubuh	Batuk Kering	Sakit Tenggorokan	Kantuk	Riwayat Perjalanan
1	55	1	1	1	1	0
2	52	1	0	0	0	1
3	53	1	1	1	1	0
4						•••
69	30	1	0	0	1	0

Gejala Tinggi: Ada beberapa gejala yang paling banyak muncul dari data pasien covid-19 yang telah di klasifikasi, yaitu suhu tubuh ≥ 39 C, batuk

Vol. 2, No. 1, February 2021 e-ISSN: xxxx-xxxx

https://journal-computing.org/index.php/journal-sea/index

kering, sakit tenggorokan, permasalahan dalam pernapasan, memiliki riwayat perjalanan, dapat dilihat pada table 6 Sekitar 0,16% pasien covid-19 merasakan seluruh gejala tersebut.

Table 6 Dari Data Klasifikasi Gejala Tinggi

No	Umur	Suhu Tubuh	Batuk Kering	Sakit Tenggorokan	Gangguan Pernapasan	Riwayat Perjalanan
1	40	1	0	0	0	1
2	33	1	0	1	1	1
3	51	1	1	1	1	0
4						
25	20	1	0	1	1	1

#### 4. KESIMPULAN

Berdasarkan hasil dan pembahasan dari dataset yang telah diuji menggunakan software rapidminer menunjukan nilai absolute count 25,98% untuk tingkat rendah, 54,33% untuk tingkat sedang, dan 19,69% untuk tingkat tinggi. Algoritma Neural Network mempunyai nilai akurasi tertinggi sebesar 73,68%, Random Forest memiliki nilai akurasi sebesar 68,42%, Naive Bayes mendapatkan nilai akurasi sebesar 65,38%, sedangkan Algoritma K-Nearest Neighbor memiliki nilai akurasi paling rendah yaitu 57,89%.

#### DAFTAR PUSTAKA

- [1] Tim Kerja Kementerian Dalam Negeri, D. S. "Pedoman Umum Menghadapi Pandemi Covid-19 bagi Pemerintah Daerah" *Jakarta, Indonesia: Kementerian Dalam Negeri RI* 2020, [Online] Available.
- [2] N. Sholikah. P.S "Kesiapsiagaan Indonesia Menghadapi Potensi Penyebaran Corona Virus Disease" *PPBK DPR RI*, vol. XII, No. 3, pp 13-18, 2020,
- [3] Abdillah, L. (2020). Stigma Terhadap Orang Positif COVID-19 (Stigma on Positive People COVID-19). *Pandemik COVID-19: Antara Persoalan Dan Refleksi Di Indonesia, Forthcoming*.
- [4] JHU CSSE COVID-19 Data, desember 2020. https://news.google.com/covid19/map?hl=id&mid=%2Fm%2F02j71&gl =ID&ceid=ID%3Aid
- [5] Siagian, T. H. (2020). Mencari kelompok berisiko tinggi terinfeksi virus corona dengan discourse network analysis. *Jurnal Kebijakan Kesehatan Indonesia: JKKI*, *9*(2), 98-106.
- [6] Buana, D. R. (2020). Analisis perilaku masyarakat indonesia dalam menghadapi pandemi virus corona (Covid-19) dan kiat menjaga kesejahteraan jiwa. *Salam: Jurnal Sosial dan Budaya Syar-i*, 7(3), 217-226

Suci Anggraini, Muhamad Akbar, Alex Wijaya, Hadi Syaputra, Muhammad Sobri 67

## https://journal-computing.org/index.php/journal-sea/index

- [7] Indrayanti, I., Sugianti, D., & Al Karomi, A. (2017). Optimasi Parameter K Pada Algoritma K-Nearest Neighbour Untuk Klasifikasi Penyakit Diabetes Mellitus. *Prosiding SNATIF*, 823-829.
- [8] Suryanegara, G. A. B., & Purbolaksono, M. D. (2021). Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(1), 114-122.
- [9] Rifai, B. (2013). Algoritma Neural Network Untuk Prediksi Penyakit Jantung. *Jurnal Techno Nusa Mandiri*, 10(1), 1-9.
- [10] Imandasari, T., Irawan, E., Windarto, A. P., & Wanto, A. (2019, September). Algoritma Naive Bayes Dalam Klasifikasi Lokasi Pembangunan Sumber Air. In *Prosiding Seminar Nasional Riset Information Science (SENARIS)* (Vol. 1, pp. 750-761).
- [11] Nofitri, R., & Irawati, N. (2019). INTEGRASI METODE NEIVE BAYES DAN SOFTWARE RAPIDMINER DALAM ANALISIS HASIL USAHA PERUSAHAAN DAGANG. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, *6*(1), 35-42.
- [12] Dinata, R. K., Akbar, H., & Hasdyna, N. (2020). Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus. *ILKOM Jurnal Ilmiah*, *12*(2), 104-111.
- [13] Ayudhitama, A. P., & Pujianto, U. (2020). Analisa 4 Algoritma Dalam Klasifikasi Liver Menggunakan Rapidminer. *Jurnal Informatika Polinema*, *6*(2), 1-9.
- [14] Siddik, M., Hendri, H., Putri, R. N., Desnelita, Y., & Gustientiedina, G. (2020). Klasifikasi Kepuasan Mahasiswa Terhadap Pelayanan Perguruan Tinggi Menggunakan Algoritma Naïve Bayes. INTECOMS: Journal of Information Technology and Computer Science, 3(2), 162-166.
- [15] Ciputra, A., Rachmawanto, E. H., & Susanto, A. (2018). Klasifikasi Tingkat Kematangan Buah Apel Manalagi dengan Algoritma Naive Bayes dan Ekstraksi Fitur Citra Digital. Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer, 9(1), 465-472.
- [16] Zailani, A. U., & Hanun, N. L. (2020). Penerapan Algoritma Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera. *Infotech: Journal of Technology Information*, 6(1), 7-14.
- [17] Religia, Y., Nugroho, A., & Hadikristanto, W. (2021). Analisis Perbandingan Algoritma Optimasi pada Random Forest untuk Klasifikasi Data Bank Marketing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(1), 187-192.
- [18] Rifai, B. (2013). Algoritma Neural Network Untuk Prediksi Penyakit Jantung. *Jurnal Techno Nusa Mandiri*, 10(1), 1-9.
- [19] Hastuti, K. (2012). Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif. *Semantik*, 2(1).