

# Implementasi Algoritma Naïve Bayes Classifier (NBC) untuk Klasifikasi Penyakit Ginjal Kronik

Qurotul A'yuniyah\*, Ena Tasia, Nanda Nazira, Pangeran Fadillah Pratama, Muhammad Ridho Anugrah, Jeni Adhiva, Mustakim

Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: <sup>1</sup>\*qurotulyuyun510@gmail.com, <sup>2</sup>enatasia6696@gmail.com, <sup>3</sup>nandanazira0211@gmail.com, <sup>4</sup>pangeranfadhilah1@gmail.com, <sup>5</sup>mridhoanugrah145@gmail.com, <sup>6</sup>jeniadhiva@gmail.com, <sup>7</sup>mustakim@uin-suska.ac.id

Email Penulis Korespondensi: qurotulyuyun510@gmail.com

Submitted: 02/09/2022; Accepted: 26/09/2022; Published: 30/09/2022

**Abstrak**—Penyakit degenerative merupakan penyakit tidak menular yang timbul akibat buruknya pola hidup yang tidak sehat, sehingga dapat menurunkan kualitas fisik dan mental penderitanya. Penyakit Ginjal Kronik (PGK) merupakan penyakit degenerative yang termasuk kedalam kategori 10 besar penyebab kematian di dunia menurut World Health Organization (WHO). Pada penelitian ini telah menggunakan data PGK dengan atribut usia, tekanan darah, berat badan, kadar albumin, kadar gula, sel darah merah, sel nanah, gumpalan sel nanah, bakteri, kadar gula dalam darah, kadar urea dalam darah, serum kreatinin, natrium, magnesium, hemoglobin, volume yang ditempati darah merah, indikasi hipertensi, indikasi diabetes mellitus, indikasi jantung koroner, nafsu makan, indikasi pembekakan pada betis atau kaki, dan indikasi anemia. Maka dilakukan klasifikasi pada data PGK dengan implementasi algoritma Naïve Bayes Classifier (NBC) yang unggul dan menghasilkan tingkat akurasi yang tinggi. Hasil klasifikasi dengan tools RapidMiner yang dilakukan dengan penerapan algoritma NBC maka diperoleh nilai akurasi 96.43%, dan hasil rata-rata recall 93.18%, rata-rata precision 93.02% dan AUC 93.2%. Sehingga disimpulkan bahwa kinerja NBC dalam mengklasifikasi data PGK kronik sangat baik.

**Kata Kunci:** Data Mining; Klasifikasi; NBC; Penyakit Ginjal Kronik; RapidMiner

**Abstract**—Degenerative disease is a non-communicable disease that arises from an unhealthy lifestyle, so that it can reduce the physical and mental quality of the sufferer. Chronic Kidney Disease (CDK) is a degenerative disease that is included in the world's top 10 causes of death according to the World Health Organization (WHO). This study used CDK data with attributes of age, blood pressure, weight, albumin levels, sugar levels, red blood cells, pus cells, pus cell clots, bacteria, blood sugar levels, blood urea levels, creatinine serum, sodium, magnesium, hemoglobin, the volume occupied by red blood, indications of hypertension, indications of diabetes mellitus, indications of coronary heart disease, appetite, indications of swelling in the calves or feet, and indications of anemia. Therefore, the classification of kidney disease data is carried out with the implementation of the superior Naïve Bayes Classifier (NBC) algorithm and produces a high level of accuracy. The classification results using the RapidMiner tools carried out by the application of the NBC algorithm, the accuracy value is 96.43%, the average recall is 93.18%, the average precision is 93.02%, and the AUC is 93.2%. so it can be concluded that the performance of NBC in classifying chronic kidney disease data is excellent.

**Keywords:** Classification; Chronic Kidney Disease; Data Mining; NBC; RapidMiner

## 1. PENDAHULUAN

Pola hidup yang sehat adalah harapan semua orang, namun kenyataannya banyak orang yang menerapkan pola hidup yang tidak sehat. Pola hidup yang tidak sehat ini dapat mengakibatkan timbulnya berbagai macam penyakit tidak menular atau penyakit *degenerative* yang dapat mengakibatkan penurunan kualitas hidup seseorang [1]. Dengan adanya kualitas hidup membuat seseorang dapat menikmati kehidupannya, kualitas hidup yang dimiliki seseorang sangat menyangkut pada kesehatan fisik dan mental. Menurut *Sample Registration Survey (SRS)* yang diteliti oleh badan pengembangan kesehatan kementerian kesehatan dan badan penelitian menyatakan bahwa penyakit *degenerative* merupakan penyakit yang mengakibatkan kematian terbanyak, salah satunya Penyakit Ginjal Kronik (PGK) [2]. Ginjal merupakan organ yang begitu penting dan memiliki fungsi yang sangat besar yaitu untuk menjaga kesehatan komposisi darah dengan cara menghindari penumpukan kotoran, serta berfungsi untuk menjaga keseimbangan cairan yang ada didalam tubuh manusia, juga untuk menjaga produksi enzim dan hormon yang akan membantu mengendalikan tekanan darah dalam tubuh dan akan menjaga tulang tetap sehat dan kuat [3].

PGK adalah penyakit yang mana prevalensi dan insidennya selalu mengalami peningkatan dan pengobatan dengan biaya terbesar dipenjuru dunia. *World Health Organization (WHO)* melaporkan bahwa PGK merupakan penyakit yang termasuk kedalam kategori 10 besar penyebab kematian di dunia. Menurut penjelasan Kementerian Kesehatan RI pada tahun 2013, jumlah penduduk yang telah mengalami PGK di Indonesia saja mencapai 499.800 penduduk [4]. Dengan banyaknya jumlah manusia yang terkena PGK, maka dibutuhkan klasifikasi untuk memprediksi PGK terhadap gejala yang menuju kepada indikator terkena PGK, maka digunakan sebuah metode dalam melakukan pengolahan data yaitu *data mining*.

*Data mining* merupakan sebuah proses yang dilakukan untuk pengambilan keputusan, juga untuk menemukan relasi antar data yang sebelumnya tidak diketahui siapa pengguna data tersebut dan proses tersebut dilakukan dengan menggunakan cara yang mudah dipahami [5]. *Data mining* juga terbagi dalam beberapa bagian diantaranya: *estimation, prediction, classification, clustering dan association*. Klasifikasi merupakan salah satu

proses yang digunakan untuk menentukan model yang membedakan kelas-kelas terhadap data. Proses klasifikasi dalam penelitian ini yaitu menggunakan metode algoritma NBC yang merupakan klasifikasi statistik yang berfungsi dapat melakukan prediksi terhadap probabilitas dalam sebuah *class* [6].

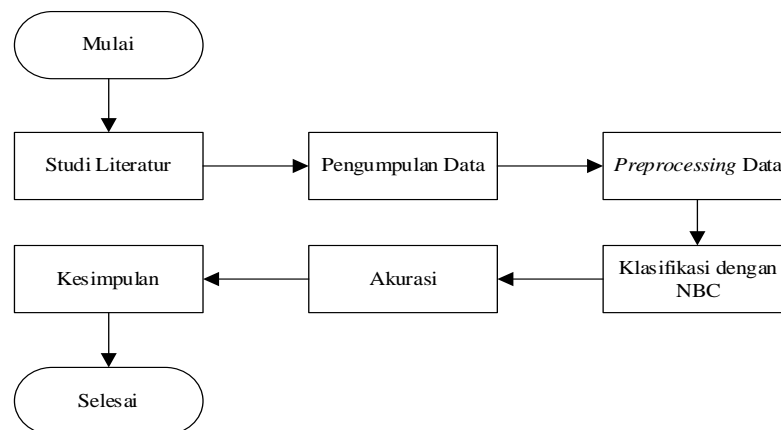
Penelitian yang dilakukan oleh Mathoril Hudha, dkk, pada tahun 2022, tentang Analisis Sentimen Pengguna *Youtube* dengan Metode NBC, hasil dari penelitian ini mengenai klasifikasi komentar dengan otomatis menggunakan bahasa pemrograman *python*, maka didapatkan nilai akurasi menggunakan algoritma NBC yaitu sebesar 90.36% [7]. Menurut penelitian terdahulu yang dilakukan oleh Astia Weni Syaputri, dkk, pada tahun 2020 tentang *NBC for Classification of Student Major's Specialization*, berdasarkan hasil dari penelitian yang telah dilakukan, bahwa spesialisasi jurusan menggunakan algoritma NBC dapat menghasilkan tingkat akurasi yang tinggi dari pemilihan jurusan, maka diperoleh akurasi sebesar 96,19% [8].

Penelitian terdahulu pada tahun 2019 yang dilakukan oleh Dudi Irawan, dkk, Tentang Analisis Penerapan Algoritma NBC Untuk Klasifikasi PGK, dari hasil penelitian tersebut saat menggunakan nilai 4 sampai 10 pada k, hasil klasifikasi dari 158 data maka didapat nilai akurasi yang tinggi adalah 100% dan nilai akurasi yang terendah adalah 99% [9]. Sedangkan penelitian yang telah dilakukan oleh Sri Widaningsih pada tahun 2019 tentang Perbandingan Metode C4.5, NBC, KNN, SVM untuk Prediksi Nilai dan Waktu Kelulusan Mahasiswa, maka hasil dari penelitian tersebut dengan menggunakan perbandingan empat algoritma dapat diketahui bahwa algoritma terbaik untuk melakukan prediksi terhadap kelulusan mahasiswa yang tepat waktu dengan nilai IPK > 3 adalah algoritma NBC dengan nilai akurasi 76.79%, dan *error* 23.17% , sedangkan algoritma SVM nilai akurasinya 74.04%, *error* 25.96%, algoritma KNN k=3 nilai akurasinya 68,05%, *error* 31,97% dan terakhir algoritma C4.5 nilai akurasinya 75.96%, *error* 24.03% [10].

Berdasarkan hasil pembahasan yang telah dipaparkan pada pendahuluan maka akan dilakukan klasifikasi terhadap pasien yang mengidap PGK. Oleh karena itu pada penelitian ini mengangkat judul yaitu tentang Implementasi Algoritma *Naïve Bayes Classifier* (NBC) untuk Klasifikasi Penyakit Ginjal Kronik.

## 2. METODOLOGI PENELITIAN

Penelitian ini menerapkan metode riset eksperimental yang terdiri atas tahapan (1) studi literatur, (2) pengumpulan data, (3) *preprocessing data*, (4) Klasifikasi dengan NBC, (5) pengujian akurasi, dan (6) penarikan kesimpulan. Secara rinci ditunjukkan pada Gambar 1.



**Gambar 1.** Metodologi Penelitian

Studi literatur dilakukan membaca dokumentasi, modul, artikel ilmiah dan sumber informasi lain yang berhubungan dengan penelitian. Data dikumpulkan dari *UCI Machine Learning Repository data set Chronic\_Kidney\_Disease Data Set*.

### 2.1 Penyakit Ginjal Kronik (PGK)

PGK merupakan perubahan fisiologis yang disebabkan oleh berbagai penyebab, sehingga masalah ini dapat menyebabkan penurunan fungsi ginjal yang *irreversible* dan progresif [11]. Resiko penyakit ini dapat menyebabkan kematian bagi penderitanya serta dapat memicu datangnya penyakit jantung. Karena ginjal yang tidak normal ini akan menimbulkan ciri-ciri seseorang terindikasi penyakit jantung. Saat ini banyak masyarakat yang tidak menjaga gaya hidup yang sehat. PGK apabila tidak ditangani dengan cepat akan menimbulkan masalah yang sangat parah [12].

### 2.2 Data Mining

*Data mining* merupakan metode atau teknik tertentu dalam data untuk mencari informasi atau pola yang menarik [13]. *Data mining* adalah istilah yang digunakan untuk mencari atau menemukan pengetahuan pada *database* [14]. *Data mining* juga dapat dikatakan sebagai kumpulan data yang tidak diperkirakan sebelumnya serta ringkas data

yang dilakukan dengan cara-cara sebelumnya yang dapat memberikan manfaat kepada pemilik data serta dapat digunakan [15].

### 2.3 Klasifikasi

Klasifikasi merupakan objek-objek yang diletakan sebelumnya dari sebuah kategori yang telah ditetapkan. Klasifikasi digunakan untuk prediksi *class* pada tabel tertentu, yakni data diklasifikasikan dari *data training* serta tabel *class*, atribut tertentu pada suatu klasifikasi dapat digunakan pula pada klasifikasi data yang baru [16]. Klasifikasi data dapat digunakan untuk pemrosesan dalam menentukan atau menjelaskan model serta fungsi dan pembeda konsepnya kelas data [17].

### 2.4 Naïve Bayes Classifier (NBC)

NBC adalah alat yang dapat digunakan untuk mengambil keputusan dari sebuah informasi [18]. Cara ini telah banyak digunakan terutama yang berkaitan dengan mendiagnosa terkait probabilitas secara statistik dan kemungkinan gejala dan penyakit terkait. Adapun keunggulan dari algoritma NBC ini yaitu metodenya yang sederhana akan tetapi memiliki hasil yang baik untuk pengklasifikasian [19]. Pengklasifikasian data dengan menggunakan algoritma ini sangat cepat dan tingginya tingkat akurasi [20]. Dalam penerapan algoritma NBC, digunakan persamaan 1[21].

$$p(W_i|C_j) = \frac{N_{cw} + 1}{N_c + V} \tag{1}$$

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan Data

Penelitian ini menggunakan pengumpulan data penderita PGK dari *UCI Machine Learning Repository*. Data menunjukkan bahwa atribut yang mempengaruhi PGK adalah usia, tekanan darah, gravitasi spesifik, kadar albumin, kadar gula, sel darah merah, sel nanah, gumpalan sel nanah, bakteri, kadar gula dalam darah, kadar urea dalam darah, serum kreatinin, natrium, magnesium, hemoglobin, volume yang ditempati sel darah merah, indikasi hipertensi, indikasi diabetes mellitus, indikasi jantung koroner, nafsu makan, indikasi pembengkakan di pada betis atau kaki, dan indikasi anemia. PGK adalah penyakit tahunan yang diderita akibat buruknya pola hidup, oleh karena itu pada penelitian ini dilakukan klasifikasi atribut data untuk menentukan nilai akurasi dalam mengidentifikasi PGK.

**Tabel 1.** Pengumpulan Data

| No. | Usia | Tekanan Darah (mm/Hg) | Gravitasi Spesifik (G) | Albumin | Gula | Sel Darah Merah | Sel Nanah | ... | Kemungkinan Ginjal Kronik |
|-----|------|-----------------------|------------------------|---------|------|-----------------|-----------|-----|---------------------------|
| 1   | 48   | 80                    | 1.020                  | 1       | 0    | Normal          | Normal    | ... | Ya                        |
| 2   | 62   | 80                    | 1.010                  | 2       | 3    | Normal          | Normal    | ... | Ya                        |
| 3   | 48   | 70                    | 1.005                  | 4       | 0    | Normal          | Abnormal  | ... | Ya                        |
| 4   | 51   | 80                    | 1.010                  | 2       | 0    | Normal          | Normal    | ... | Ya                        |
| ... | ...  | ...                   | ...                    | ...     | ...  | ...             | ...       | ... | ...                       |
| 400 | 58   | 80                    | 1.025                  | 0       | 0    | Normal          | Normal    | ... | Tidak                     |

### 3.2 Preprocessing Data

Proses selanjutnya dilakukan *data cleaning*, *Data cleaning* merupakan aktifitas membersihkan data non-relevan dan data *noise* yang dapat berupa *record* yang hilang, invalid, atau salah ketik. Pembersihan dilakukan dengan membuang data seperti yang dideskripsikan sebelumnya [22]. *Data cleaning* berpengaruh terhadap *data mining* akibat pengurangan jumlah dan kompleksitas data. Tahap *preprocessing* berawal dari sebanyak 400 *record* data. Setelah *cleaning*, jumlah data berkurang menjadi 280 *record* data.

**Tabel 2.** *Data Cleaning*

| No. | Usia | Tekanan Darah (mm/Hg) | Gravitasi Spesifik (G) | Albumin | Gula | Sel Darah Merah | Sel Nanah | ... | Kemungkinan Ginjal Kronik |
|-----|------|-----------------------|------------------------|---------|------|-----------------|-----------|-----|---------------------------|
| 1   | 48   | 80                    | 1.020                  | 1       | 0    | Normal          | Normal    | ... | Ya                        |
| 2   | 62   | 80                    | 1.010                  | 2       | 3    | Normal          | Normal    | ... | Ya                        |
| 3   | 48   | 70                    | 1.005                  | 4       | 0    | Normal          | Abnormal  | ... | Ya                        |
| 4   | 51   | 80                    | 1.010                  | 2       | 0    | Normal          | Normal    | ... | Ya                        |
| ... | ...  | ...                   | ...                    | ...     | ...  | ...             | ...       | ... | ...                       |

| No. | Usia | Tekanan Darah (mm/Hg) | Gravitasi Spesifik (G) | Albumin | Gula | Sel Darah Merah | Sel Nanah | ... | Kemungkinan Ginjal Kronik |
|-----|------|-----------------------|------------------------|---------|------|-----------------|-----------|-----|---------------------------|
| 280 | 58   | 80                    | 1.025                  | 0       | 0    | Normal          | Normal    | ... | Tidak                     |

Data yang telah melewati proses *cleaning* selanjutnya masuk ke tahap transformasi data, mengubah data menjadi lebih sederhana sehingga bentuknya sesuai dengan proses *Data mining*. Bentuk yang sesuai untuk diolah dengan menggunakan *Tools RapidMiner* adalah hasil dari proses transformasi data.

**Tabel 3.** Data Transformasi

| No. | Usia | Tekanan Darah (mm/Hg) | Gravitasi Spesifik (G) | Albumin | Gula | Sel Darah Merah | Sel Nanah | ... | Kemungkinan Ginjal Kronik |
|-----|------|-----------------------|------------------------|---------|------|-----------------|-----------|-----|---------------------------|
| 1   | 48   | 80                    | 1.020                  | 1       | 0    | 1               | 1         | ... | Ya                        |
| 2   | 62   | 80                    | 1.010                  | 2       | 3    | 1               | 1         | ... | Ya                        |
| 3   | 48   | 70                    | 1.005                  | 4       | 0    | 1               | 2         | ... | Ya                        |
| 4   | 51   | 80                    | 1.010                  | 2       | 0    | 1               | 1         | ... | Ya                        |
| ... | ...  | ...                   | ...                    | ...     | ...  | ...             | ...       | ... | ...                       |
| 280 | 58   | 80                    | 1.025                  | 0       | 0    | 1               | 1         | ... | Tidak                     |

### 3.3 Implementasi Algoritma NBC

Pada implementasi dengan metode NBC, dataset sebanyak yang telah ditransformasi dipisah dengan cara *holdout*, sehingga menghasilkan data *training* dan data *testing* dengan perbandingan 7:3. Melalui pembagian data, maka dari 280 *record* diperoleh sebanyak 196 data *training* dan 84 data *testing*. Implementasi NBC dilakukan dengan menggunakan *tools RapidMiner*. Selanjutnya pada dataset pasien ginjal kronik terdapat dua kelas yaitu “Ya Ginjal Kronik” dan “Tidak Ginjal Kronik”.

Selanjutnya akan menghitung performa klasifikasi dengan menggunakan data pengujian sebanyak 84 data uji, yang berstatus jika ya sebanyak 41 data, jika tidak tapi ya sebanyak 3 data, dan jika tidak sebanyak 40 data. Adapun hasil pada tahap pengujian akan dinyatakan dalam bentuk nilai *Precision*, *Recall*, dan *Accuracy* yang akan dihasilkan dalam bentuk *Confussion Matrix*.

**Tabel 4.** *Confussion matrix*

|              | True Ya | True Tidak | Class Precision |
|--------------|---------|------------|-----------------|
| Pred. Ya     | 41      | 0          | 100.00%         |
| Pred. Tidak  | 3       | 40         | 93.02%          |
| Class Recall | 93.18%  | 100.00%    |                 |

*Confussion Matrix* yang terdapat pada Tabel 4, menunjukkan sejumlah data *training* dan data *testing* yang merupakan hasil prediksi pada setiap kelas, presisi, dan *recall*. *Confussion Matrix* juga memberi informasi hasil nilai *true positive*, dan nilai *true negative*. Hasil pengujian *Confussion Matrix* dari dua kelas status pasien yang dapat dianalisis dengan bentuk hasil rata-rata *precision*, rata-rata *recall*, *accuracy*, dan AUC.

**Tabel 5.** Hasil data Pengujian

| Parameter        | Nilai  |
|------------------|--------|
| <i>Precision</i> | 93.02% |
| <i>Recall</i>    | 93.18% |
| <i>Accuracy</i>  | 96.43% |
| AUC              | 93.2%  |

Berdasarkan dari tabel 5 hasil klasifikasi menggunakan *tools RapidMiner* dengan algoritma NBC diperoleh nilai akurasi sebesar 96.43%.

## 4. KESIMPULAN

Data PGK yang dikumpulkan dari *UCI Machine Learning Repository* sebanyak 400 *record*, melalui *preprocessing* menjadi 280 data berupa 196 data *training* dan 84 data *testing*. Klasifikasi yang dilakukan pada penelitian ini menggunakan *tools RapidMiner* dengan algoritma NBC, validasi dalam menguji nilai akurasi maka diperoleh hasil bahwa algoritma NBC memiliki akurasi yang tinggi dengan presentase 96.43%.

## REFERENCES

- [1] N. D. Lika, “PENERAPAN ALGORITMA NBC UNTUK KLASIFIKASI TINGKAT RESIKO PENYAKIT DIABETES MELLITUS,” *J. Chem. Inf. Model.*, vol. 21, no. 1, pp. 1–9, 2020.
- [2] Suwanti, Taufikurrahman, M. I. Rosyidi, and A. Wakhid, “Kualitas Hidup Pasien Gagal Ginjal Kronik Yang Menjalani

- Terapi Hemodialisa,” *J. Keperawatan Muhammadiyah Bengkulu*, vol. 9, no. 2, pp. 1–9, 2021, doi: 10.36085/jkmb.v9i2.1711.
- [3] T. Arifin and D. Ariesta, “Prediksi Penyakit Ginjal Kronis Menggunakan Algoritma Naive Bayes Classifier Berbasis Particle Swarm Optimization,” *J. Tekno Insentif*, vol. 13, no. 1, pp. 26–30, 2019, doi: 10.36787/jti.v13i1.97.
- [4] I. Yulianti, R. Amegia Saputra, M. Sukrisno Mardiyanto, and A. Rahmawati, “Optimasi Akurasi Algoritma C4.5 Berbasis Particle Swarm Optimization dengan Teknik Bagging pada Prediksi Penyakit Ginjal Kronis Optimization of C4.5 Algorithm Based On Particle Swarm Optimization with Bagging Technique on Prediction of Chronic Kidney Dise,” *Techno.COM*, vol. 19, no. 4, pp. 411–421, 2020.
- [5] I. Kamila, U. Khairunnisa, and M. Mustakim, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau,” *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 5, no. 1, p. 119, 2019, doi: 10.24014/rmsi.v5i1.7381.
- [6] E. T. Lestari and J. Adhiva, “Implementation Naive Bayes Classifier Algorithm and K-Nearest Neighbor For Obesity Nutritional Status of Children with Disabilities Implementasi Algoritma Naive Bayes Classifier dan K-Nearest Neighbor Untuk Klasifikasi Status Gizi Obesitas Anak Disabilitas,” pp. 1–11, 2022.
- [7] M. Hudha, ---/Endang Supriyati, and T. Listyorini, “Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamenantiterawan Dengan Metode Naive Bayes Classifier,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 1, pp. 1–6, 2022, doi: 10.33387/jiko.v5i1.3376.
- [8] A. W. Syaputri, E. Irwandi, and Mustakim, “Naive Bayes Alghorithm for Classification of Student Major’s Specialization,” vol. 1, no. 1, pp. 1–5, 2020.
- [9] D. Irawan, H. Oktavianto, M. K. Anam, T. Informatika, and U. M. Jember, “Analisis Penerapan Algoritma Naive Bayes,” *JASIE “Jurnal Apl. Sist. Inf. Dan Elektron.”*, vol. 1, no. 2, pp. 127–134, 2019.
- [10] S. Widaningsih, “Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naive Bayes, Knn Dan Svm,” *J. Tekno Insentif*, vol. 13, no. 1, pp. 16–25, 2019, doi: 10.36787/jti.v13i1.78.
- [11] A. Abdullah, D. Salwani, A. B. Khairi, M. Muhsin, and M. Syukri, “Puasa ramadhan dan pengaruhnya terhadap progresifitas penyakit ginjal kronik,” *J. Kedokt. Syiah Kuala*, vol. 21, no. 3, p. 2021, 2021, doi: 10.24815/jks.v21i3.23754.
- [12] A. Sembiring, S. Andryana, and A. Gunaryati, “Sistem Pakar Berbasis Mobile Untuk Diagnosis Penyakit Ginjal Menggunakan Metode Forward Chaining,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 139–148, 2021, doi: 10.29100/jipi.v6i1.1932.
- [13] N. Nuraeni, “Klasifikasi Data Mining untuk Prediksi Potensi Nasabah dalam Membuat Deposito Berjangka,” *J. Ilm. Intech Inf. Technol. J. ....*, vol. 3, no. 01, pp. 65–74, 2021.
- [14] B. A. Candra Permana and I. K. Dewi Patwari, “Komparasi Metode Klasifikasi Data Mining Decision Tree dan Naive Bayes Untuk Prediksi Penyakit Diabetes,” *Infotek J. Inform. dan Teknol.*, vol. 4, no. 1, pp. 63–69, 2021, doi: 10.29408/jit.v4i1.2994.
- [15] J. Adhiva, S. A. Putri, and S. G. Setyorini, “Prediksi Hasil Produksi Kelapa Sawit Menggunakan Model Regresi Pada PT .Perkebunan Nusantara V,” *Semin. Nas. Teknol. Informasi, Komun. dan Ind.*, pp. 155–162, 2020.
- [16] A. Tarigan, E. Wahyudi, and J. Adhiva, “Klasifikasi Status Kesejahteraan Rumah Tangga di Kabupaten Siak Menggunakan Algoritma Naive Bayes Classifier,” no. November, pp. 187–196, 2019.
- [17] A. Hamdani, Mustakim, and I. Kamila, “Klasifikasi Dokumen Tugas Akhir Berbasis Text Mining menggunakan Metode Naive Bayes Classifier dan K-Nearest Neighbor,” *Semin. Nas. Teknol. Informasi, Komun. dan Ind. II*, no. November, pp. 178–186, 2019.
- [18] T. A. Putra and P. A. . Purnama, “Perancangan Sistem Pakar untuk Mendiagnosa Penyakit Toksoplasma pada Wanita Menggunakan Metode Bayes dengan Bahasa Pemrograman PHP dan Database MySQL,” *Sink. J. dan Penelit. Tek. Inform.*, vol. 3, no. 1, pp. 120–129, 2018.
- [19] firman Tempola, “Implementasi Metode Naive Bayes Untuk Memprediksi Resiko Penyakit Jantung,” *Patria Artha Technol. J.*, vol. 4, no. 2, pp. 66–70, 2020, doi: 10.33857/patj.v4i2.351.
- [20] A. Ali, A. Khairan, F. Tempola, and A. Fuad, “Application Of Naive Bayes to Predict the Potential of Rain in Ternate City,” *E3S Web Conf.*, vol. 328, p. 04011, 2021, doi: 10.1051/e3sconf/202132804011.
- [21] A. Harun and D. P. Ananda, “Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naive Bayes dan Decision Tree,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 58–63, 2021.
- [22] R. Rustiyan and M. Mustakim, “Penerapan Algoritma Fuzzy C Means untuk Analisis Permasalahan Simpanan Wajib Anggota Koperasi,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 171, 2018, doi: 10.25126/jtiik.201852605.