

Systematic Literature Review: Perbandingan Algoritma Klasifikasi

Prayugo Pangestu¹, Rice Novita², Mustakim³

^{1,2,3} Sistem Informasi, UIN Sultan Syarif Kasim Riau, Pekanbaru, Riau, Indonesia

Email: 11950314979@students.uin-suska.ac.id¹, rice.novita@uin-suska.ac.id², mustakim@uin-suska.ac.id³

Abstract – Over time, many data mining methods have been created and suggested to help in making decisions. Due to limited resources, this article only provides a systematic literature review in comparing the performance of the Naïve Bayes, Decision Tree, Neural Network, Random Forest, and Support Vector Machine methods to find out which method is most effective in classifying and predicting. After conducting a literature study by taking articles from 2019 to 2023, 500 articles were obtained that used the Naive Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine methods. Because there were so many articles obtained in the initial search, inclusion and exclusion criteria were created to sort out articles that were in accordance with this research. After carrying out the inclusion and exclusion criteria process, 243 articles were obtained and it was discovered that the topic that was discussed more often was prediction, which amounted to 122 articles and the remaining 121 articles discuss classification. In the field of prediction, the method most frequently used is Random Forest with a total of 45 articles and an average accuracy rate of 91.18%, while in the field of classification, the method most frequently used is Support Vector Machine with a total of 32 articles and an average accuracy of 88.85%.

Keywords - naïve bayes, decision tree, neural network, random forest, support vector machine

Intisari – Seiring waktu, banyak metode data mining yang dibuat dan disarankan untuk membantu dalam mengambil keputusan. Karna keterbatasan dalam sumberdaya atikel ini hanya memberi kajian literatur secara sistematis dalam membandingkan performa metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine untuk mengetahui metode mana yang paling efektif digunakan dalam mengklasifikasi dan memprediksi. Setelah dilakukan studi literatur dengan mengambil artikel dari rentang waktu 2019 sampai 2023 didapatkan sebanyak 500 artikel yang menggunakan metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine. Karna artikel yang didapat dalam pencarian awal begitu banyak, dibuatlah kriteria inklusi dan eksklusi untuk memilah artikel yang memang sesuai dengan penelitian ini, setelah melakukan proses kriteria inklusi dan eksklusi, didapatkan sebanyak 243 artikel dan diketahui bahwa topik yang lebih banyak dibahas adalah prediksi, yaitu berjumlah 122 artikel dan sisanya 121 artikel membahas tentang klasifikasi. Pada bidang prediksi metode yang paling sering digunakan adalah Random Forest dengan jumlah 45 artikel dan rata rata tingkat akurasi 91,18%, sedangkan pada bidang klasifikasi metode yang paling sering digunakan adalah Support Vector Machine dengan total 33 artikel dan rata rata akurasi 88,85%.

Kata Kunci - Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine

I. PENDAHULUAN

Data mining adalah sebuah cara dalam menemukan sebuah informasi penting yang sebelumnya tidak diketahui [1]. Tahapan yang dilakukan pada proses data mining adalah, menyeleksi data, selanjutnya dilakukan *processing* agar kualitas data menjadi lebih baik, selanjutnya dilakukan transformasi, serta interpretasi, dan yang terakhir evaluasi, agar output yang dihasilkan menjadi sebuah pengetahuan baru yang bisa memberikan dampak yang lebih baik [2]. Data mining juga memiliki beberapa peran penting, diantaranya adalah klasifikasi dan prediksi [3]. Klasifikasi sendiri merupakan sebuah kegiatan dalam menilai suatu objek data

untuk mengelompokkannya kedalam kelas tertentu dari beberapa jenis kelas yang tersedia [4]. Klasifikasi mampu mereplikasi sebuah model pembangunan berdasarkan data yang didapat, dan menggunakan data replika tersebut untuk membangun data yang baru [5]. Klasifikasi bisa diasumsikan seperti kegiatan dalam mempelajari sebuah pola data sehingga didapatkan data yang lebih terperinci didalam data tersebut [6]. Sedangkan prediksi merupakan kegiatan dalam meramalkan atau memperkirakan kemungkinan-kemungkinan yang bisa saja akan terjadi [7]. Prediksi juga erat kaitannya dengan pengambilan keputusan. Didalam prediksi diperlukan sebuah proses yang digunakan dalam menentukan pilihan atau hasil akhir yang akan digunakan. Selain itu, keputusan yang diambil pun tidak bisa diambil dengan tidak sengaja atau kebetulan [8].

Beberapa tahun terakhir banyak metode data mining yang telah dibuat dan disarankan dalam melakukan proses klasifikasi dan prediksi, diantaranya Naïve bayes, decision tree, neuralnetwork, random forest dan support vactor machine. Lima metode tersebut pastinya memiliki keunggulan. Naïve bayes adalah metode yang bisa digunakan dalam berbagai jenis data (kuantitatif, kualitatif) dan data training yang perlukan relatif sedikit [9], sedangkan decision tree memiliki tingkat keakuratan yang bisa dibilang tinggi dan bisa digunakan dalam mengolah data yang memiliki multidimensi[10], neural network merupakan metode yang memiliki kemungkinan gagal yang terbilang sangat kecil karna dalam prosesnya metode ini memproses data secara bolak balik[11], selanjutnya random forest, metode ini bisa menyelesaikan suatu permasalahan dengan efisien, dan memiliki tingkat eror yang terbilang rendah[12], dan yang terakhir support vector machine metode ini tetap bisa mendapatkan hasil yang maksimal meskipun hanya memiliki sedikit data training, karna untuk melatih metode ini hanya memerlukan data yang relatif sedikit[13].

Karna semakin banyak nya pilihan metode yang bisa digunakan dalam menyelesaikan masalah klasifikasi dan prediksi, penelitian ini bertujuan untuk mendukung para pengambil keputusan dalam memilih metode klasifikasi pada data mining yang memiliki begitu banyak pilihan, serta mengetahui metode yang paling pas digunakan dalam menyelesaikan masalah di topik dan bidang tertentu dengan hasil dari perbandingan metode melalui literature riview dan menyimpulkan posisi saat ini dalam ruang lingkup data mining. Kesimpulan didalamnya mencakup topik, bidang, tahun publish, distribusi publikasi, penulis, serta jurnal publikasi. Target utama penelitian yang akan dilakukan adalah mengetahui performa metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, dan Suport Vector Machine dengan melihat rata rata akurasi dari setiap metode serta melihat metode mana yang paling sering digunakan dalam topik klasifikasi dan juga prediksi pada lima tahun terakhir, diharapkan hasil akhir dari penelitian ini bisa membantu para peneliti dalam memilih metode yang ingin digunakan dalam menyelesaikan masalah dibidang tertentu, serta membantu dalam memilih metode yang ingin digunakan dalam mengklasifikasi dan memprediksi

II. SIGNIFIKASI STUDI

A. Metode Penelitian

Saat ini tinjauan secara literature riview menjadi salah satu pilihan didalam mengolah data. Literature Riview merupakan suatu proses menafisir, mengidentifikadi, dan menilai hasil dari semua penelitian yang tersedia dalam rangka mencari jawaban pertanyaan penelitian tertentu [14]. Metode yang digunakan, gaya penulisan, dibagian ini juga terinspirasi dari [15] dan [16]. Berikut merupakan metodologi penelitian yang digunakan dalam penelitian ini:



Gambar 1. Metode Penelitian

Seperti yang ditunjukkan pada gambar 1, dalam literature rievew terdapat tiga tahap didalamnya, diantaranya adalah: merencanakan, melaksanakan dan pelaporan. Pada (langkah 1) akan dilakukan ide tifikasi kebutuhan yang ditinjau secara sistematis yang ada mengenai metode Naïve bayes, Decision Tree, Neural Network, Random Forest, dan Support Vector Machine, selanjutnya akan di kembangkan sebuah protokol pencarian tinjauan, agar penelitian yang dilakukan tidak keluar dari tujuan penelitian. (tahap 2) selanjutnya akan dicari dan dipilih berbagai artikel dari web jurnal yang sudah terindeks sinta dan scopus, setelah didapatkan artikel dengan jumlah yang diinginkan, artikel tersebut akan di pilah berdasarkan rentang inklusi dan eksklusi. Pada (langkah 3) akan dibuat sebuah rangkuman penelitian dalam bentuk laporan penelitian

B. Pertanyaan Penelitian

Research questions (RQ) / Pertanyaan penelitian berfungsi sebagai sebuah acuan atau jalur yang sengaja dibuat, agar penelitian yang dilakukan tetap berada pada jalur yang diharapkan. Hal ini dirancang berdasarkan *Population, Intervention, Comparison, Outcomes, Context* (PICOC) [17]. Berikut adalah tampilan dari PICOC dalam penelitian ini

TABEL I
KETERANGAN PICOC

Population	Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine
Intervention	Kekurangan dan kelebihan setiap metode, Performa metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine pada lima tahun terakhir (jurnal penerbit, jumlah artikel pertahun, performa pada setiap bidang ilmu),
Comparison	-
Outcomes	Performa metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine dalam bidang klasifikasi dan prediksi yang dihitung dalam tingkat akurasi
Context	Studi yang dilakukan secara literatur, dataset kecil dan besar

Selanjutnya pada tabel II akan dijelaskan pertanyaan penelitian yang dibahas dalam tinjauan literature

TABEL II
PERTANYAAN PENELITI

ID	Pertanyaan
RQ1	Jurnal mana yang paling banyak menerbitkan artikel yang menggunakan metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine?
RQ2	Bagaimana performa masing masing metode pada lima tahun terakhir dari rentang 2019-2023?
RQ3	Masalah apa yang paling banyak diteliti setiap metode pada rentang tahun 2019-2023?
RQ4	Metode yang sering dipilih untuk menyelesaikan kasus kasus yang berkaitan dengan klasifikasi?
RQ5	Metode yang sering dipilih untuk menyelesaikan kasus kasus yang berkaitan dengan prediksi?
RQ6	Dari ke 5 metode tersebut, mana yang paling efektif digunakan dalam menyelesaikan masalah klasifikasi dan prediksi?

Dari studi utama, Perbandingan metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine diperlukan kerangka kerja dan kumpulan data untuk menjawab pertanyaan RQ4 hingga RQ6 sehingga Pertanyaan RQ4 hingga RQ6 merupakan pertanyaan utama, dan pertanyaan lainnya (RQ1 hingga RQ3) membantu kita dalam mengevaluasi konteks utama penelitian dengan memberikan sedikit gambaran performa metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine.

C. Pemilihan Studi

Dalam mencari data untuk yang pertama kali didapatkan begitu banyak jurnal sehingga dibuatlah kriteria inklusi dan eksklusi agar artikel yang didapatkan sesuai dengan yang dicari

TABEL III
KRITERIA INKLUSI DAN EKSKLUSI

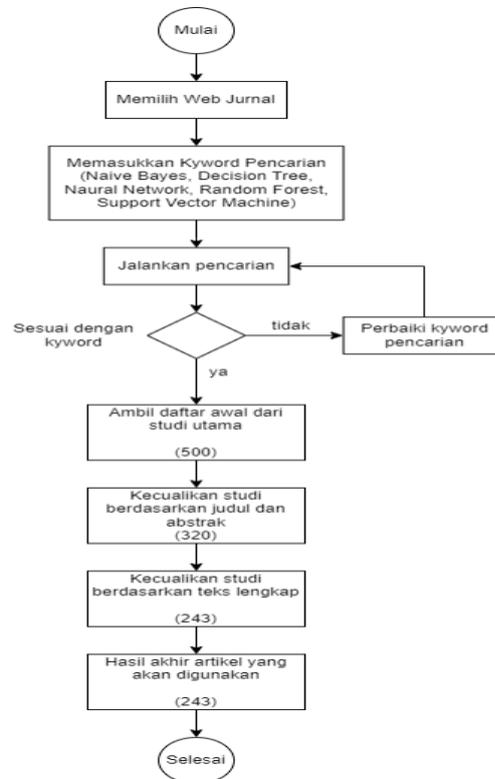
Inklusi	<ul style="list-style-type: none"> a. Studi yang dipublish dalam rentang tahun 2019 sampai 2023 b. Studi yang terindeks Minimal Sinta 3 dan Scopus Q2 c. Dataset premier maupun skunder d. Studi yang memiliki hasil akhir akurasi dari metode yang digunakan
Eksklusi	<ul style="list-style-type: none"> a. Studi yang menggunakan Bahasa Indonesia dan Inggris b. Studi yang menyertakan hasil eksperimen dalam penelitiannya c. Studi yang menggunakan Metode naive bayes, neural network, decision tree, random forest, support vector machine

D. Analisa dan Hasil

Peneliti juga memanfaatkan Mendeley (<http://mendeley.com>) sebagai sebuah tempat dalam menyimpan dan mengelola artikel-artikel yang sudah didapatkan. Proses pencarian dan jumlah artikel selama penelitian yang diidentifikasi pada setiap tahap ditunjukkan pada gambar 2. Seperti yang ditampilkan pada gambar 2, proses pemilihan artikel didalam penelitian dilakukan dalam dua langkah: judul dan abstrak dari artikel yang ingin digunakan sebagai bahan studi literature akan dilihat, selanjutnya dilihat dalam teks lengkap. Penelitian yang tidak menampilkan hasil eksperimen tidak akan dimasukkan.

Hasil akhir artikel yang terpilih untuk lanjut ke tahap mengolah data sebanyak 243 artikel. Selanjutnya, sebanyak 243 artikel tersebut dianalisis, selain kriteria inklusi dan eksklusi yang

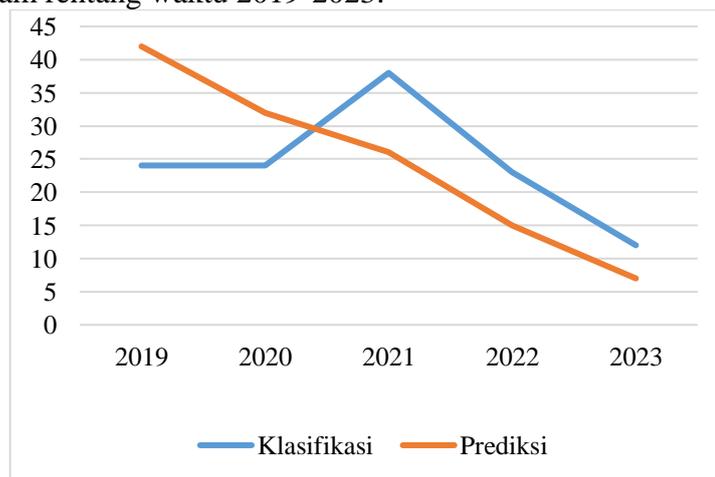
sudah dibuat sebelumnya, relevan atau tidaknya artikel dengan RQ pertanyaan peneliti, dan kemiripan antar artikel juga dipertimbangkan.



Gambar 2. Proses Pencarian Artikel

III. HASIL DAN PEMBAHASAN

Berikut merupakan grafik artikel yang membahas topik klasifikasi dan juga prediksi dari tahun ke tahun dalam rentang waktu 2019-2023:

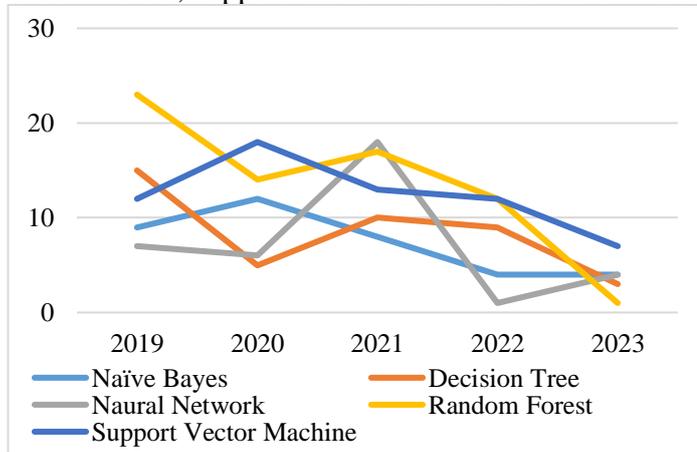


Gambar 3. Artikel yang Membahas Topik Klasifikasi dan Prediksi

Dari masing masing topik yang diteliti selama 5 tahun terakhir, terdapat selisih yang sangat tipis, yaitu hanya selisih 1 artikel. Setelah didapatkan 243 artikel yang siap untuk diolah, sebanyak 121 artikel diantaranya membahas topik klasifikasi dalam penelitiannya, sedangkan artikel yang membahas topik prediksi berjumlah 123.

Dalam mencari beberapa artikel yang menggunakan metode naïve bayes, decision tree, neural network, random forest, support vector machine guna melakukan tinjauan literature,

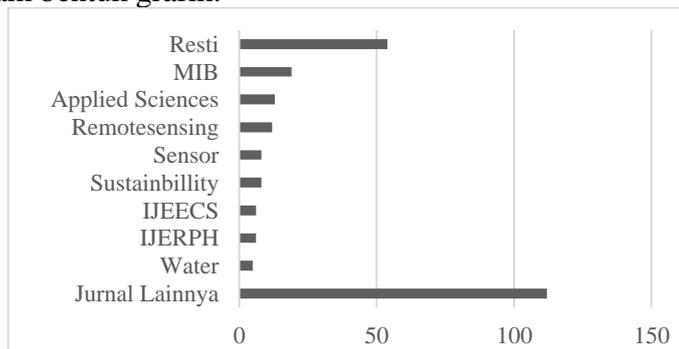
diperlukan beberapa sumber jurnal. Setelah dilakukan pencarian serta pengelolaan data didapatkan, artikel yang menggunakan metode Naïve Bayes sebanyak 37 artikel, Decision Tree 41 artikel, Neural Network 36 artikel, Random Forest 67 artikel, dan Support Vector Machine 62 artikel. Berikut ini adalah grafik yang menjelaskan penggunaan metode naïve bayes, decision tree, neural network, random forest, support vector machine dari tahun ke tahun:



Gambar 4. Penggunaan Metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine Berdasarkan Tahun

Pada grafik tersebut, bisa dilihat bahwa metode naïve bayes banyak digunakan pada tahun 2020 dengan jumlah 12 artikel, sedangkan studi menggunakan metode Decision Tree paling banyak digunakan pada tahun 2019 yaitu berjumlah 15 artikel, untuk studi menggunakan metode Neural Network paling banyak digunakan pada tahun 2021 yaitu berjumlah 18 artikel dari jumlah 36 artikel, sedangkan studi menggunakan metode Random Forest paling banyak digunakan pada tahun 2019 yaitu berjumlah 23 artikel dari jumlah 67 artikel, dan studi menggunakan metode Suport Vector Machine paling banyak digunakan pada tahun 2020 yaitu berjumlah 18 artikel dari jumlah 62 paper.

Studi menggunakan metode naïve bayes, decision tree, neural network, random forest, support vector machine sudah banyak dipublish pada beberapa jurnal besar baik nasional maupun internasional, berikut merupakan beberapa jurnal penerbit artikel yang menggunakan metode naïve bayes, decision tree, neural network, random forest, support vector machine yang akan dijabarkan dalam bentuk grafik:

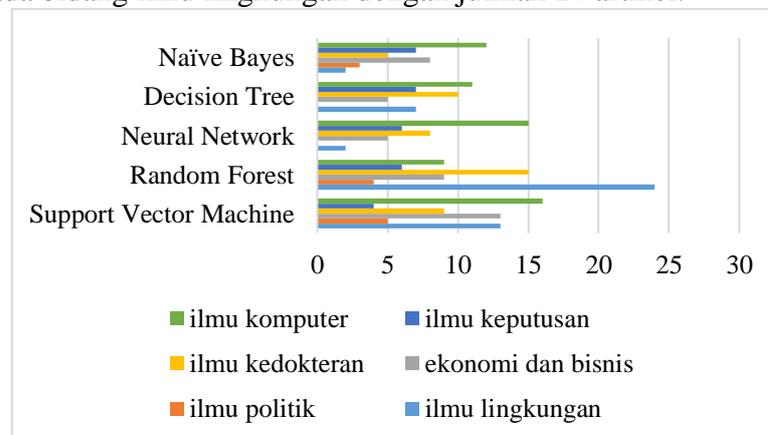


Gambar 5. Jurnal Penerbit

Pada lima tahun terakhir, metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, Support Vector Machine sering dimanfaatkan para peneliti dalam berbagai macam bidang, seperti: ilmu komputer, ilmu keputusan, ilmu kedokteran, ekonomi dan bisnis, ilmu politik, ilmu lingkungan.

Pada Gambar 5 bisa dilihat bahwa metode Naïve Bayes, Decision Tree, Neural Network, dan Support Vector Machine lebih sering dimanfaatkan pada bidang ilmu komputer dengan

Naïve Bayes 12 artikel, Decision Tree 11 artikel, Neural Network 15 artikel, Support Vector Machine 16 artikel, sedangkan metode Random Forest pada lima tahun terakhir sering dimanfaatkan pada bidang ilmu lingkungan dengan jumlah 24 artikel.

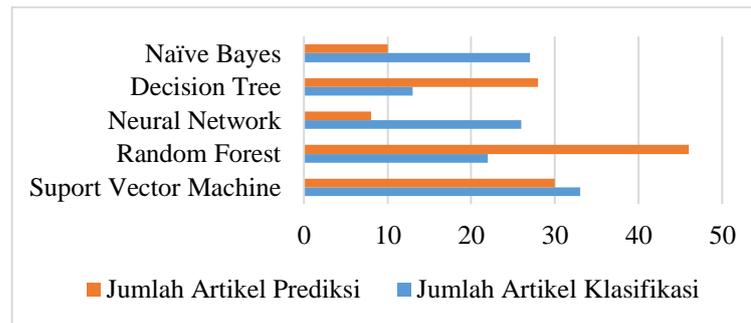


Gambar 6. Penggunaan Metode Naive Bayes, Decision Tree, Neural Network, Random Forst, Support Vector Machine Berdasarkan Bidang.

Setelah dilakukan studi literature diketahui Naïve Bayes memiliki kelebihan yaitu: bisa digunakan dalam berbagai jenis data (kuantitatif, kualitatif) dan Data training yang perlukan hanya sedikit, dan Bisa dipakai untuk mengklasifikasi jenis data biner maupun multiclass [18], tetapi Naïve bayes memiliki kelemahan dalam membuat sebuah keputusan, karna naïve bayes sangat bergantung dengan kasus terdahulu, dan hanya bisa mendeteksi data yang berupa teks, tidak dengan gambar [19]. Selanjutnya Decision Tree memiliki kelebihan tingkat keakuratannya bisa dibilang tinggi dan bisa digunakan dalam mengolah data yang memiliki multidimensi [20], tetapi Decision Tree lemah jika dilakukan perubahan atau penambahan data didalamnya, karna bisa memakan waktu yang cukup lama [21]. Neural Network memiliki kelebihan kemungkinan gagalnya terbilang sangat kecil karna dalam prosesnya metode ini memproses data secara bolak balik [22], dan memiliki kelemahan tahapan dalam memproses dengan neural network termasuk sangat panjang dan sulit dalam menginterpretasikan dikarenakan analisisnya sangat kompleks [23]. Random Forest memiliki kelebihan jika data training yang diberikan pada metode ini banyak, metode ini bisa menyelesaikan dengan efisien, dan memiliki tingkat eror yang terbilang rendah [24], karna dalam prosesnya random forest mengolah data secara acak, kita tak memiliki kontrol yang leluasa [25]. Terakhir Support Vector Machine, memiliki kelebihan tetap bisa mendapatkan hasil yang maksimal meskipun hanya memiliki sedikit data training, karna untuk melatih metode ini hanya memerlukan data yang relative sedikit [26], tetapi Support Vector Machine memiliki kelemahan sulit dalam mengaplikasikan jika himpunan / data yang diolah memiliki jumlah dimensi yang besar [27].

Gambar 7 menunjukkan, pada lima tahun terakhir 3 metode yang sering terpilih dalam menyelesaikan topik prediksi, diantaranya adalah Random Forest, Suport Vector Machine, dan Decision Tree. Random Forest dengan total artikel yang meneliti bidang prediksi sebanyak 45 artikel, selanjutnya ada Suport Vector Machine dengan jumlah artikel sebanyak 30 artikel, dan yang terakhir Decision Tree, berjumlah 26 artikel.

Sedangkan pada bidang klasifikasi, 3 metode yang sering digunakan dalam menyelesaikan masalah klasifikasi adalah Suport Vector Machine, Neive Bayes, Neural Network. Sebanyak 32 artikel menggunakan metode Support Vector Machine dalam menyelesaikan masalah klasifikasi, selanjutnya Naïve Bayes, sebanyak 27 artikel, dan yang terakhir Neural Network, dalam beberapa tahun terakhir Neural Network menjadi salah satu pilihan dalam menyelesaikan masalah yang ada didalam topik klasifikasi, dengan jumlah 26 artikel, berikut ditampilkam dalam bentuk grafik:



Gambar 7. Grafik Jumlah Artikel Klasifikasi dan Prediksi pada Setiap Metode

Pada tabel 4 menunjukkan presentase dari studi yang menggunakan metode Naïve Bayes, Decision Tree, Neural Network, Random Forest, dan Support Vector Machine pada topik tertentu dibagi dengan jumlah yang sama. Tools yang digunakan dalam mencari rata-rata presentase pada artikel ini adalah SPSS, dengan mengambil dan menjumlahkan hasil dari setiap artikel yang dilakukan studi literature dalam bentuk akurasi dan dibagi setara dengan jumlah artikel yang didapat. Bisa dilihat pada Tabel 4, pada topik klasifikasi tiga metode yang memiliki tingkat akurasi terbaik adalah metode Neural Network dengan tingkat akurasi 93,39%, selanjutnya ada Random Forest dengan tingkat akurasi 92,24%, dan yang terakhir Decision Tree dengan akurasi 91,56%. Sedangkan pada bidang prediksi, tiga metode yang memiliki tingkat akurasi terbaik adalah Neural Network dengan tingkat akurasi 93,54%, Random Forest dengan presentase akurasi 91,18%, dan yang terakhir Support Vector Machine dengan presentase akurasi mencapai 89,14%.

TABEL IV
TINGKAT AKURASI SETIAP METODE

Algoritma	Klasifikasi	Prediksi
Naïve Bayes	87,45%	77,51%
Decision Tree	91,56%	88,20%
Neural Network	93,39%	93,54%
Random Forest	92,24%	91,18%
Support Vector Machine	88,85%	89,14%

Secara keseluruhan, setiap metode data mining memiliki performa yang bisa dibilang baik, tetapi pada 5 tahun terakhir metode Support Vector Machine menjadi pilihan terfavorit dalam meneliti bidang klasifikasi, dengan jumlah artikel sebanyak 32 dari 121 artikel yang meneliti bidang klasifikasi, dengan rata rata akurasi 88.85%, sedangkan jika dilihat dari nilai akurasinya, Neural Network lah yang terbaik, yaitu 93.39% dan jumlah artikel 26 artikel. Sedangkan pada bidang prediksi, metode yang sering digunakan pada 5 tahun terakhir adalah Random Forest, dengan jumlah artikel sebanyak 45 dari 122 artikel, dan nilai rata ratanya yang sangat baik, yaitu 91.18%, tahun 2019 merupakan tahun paling banyak diterbitkannya artikel yang menggunakan metode Random Forest dengan jumlah 23 artikel dari 67 artikel. Sedangkan jika melihat dari rata rata akurasi dari kelima metode, Neural Network lah yang paling baik, yaitu 93,54% dari 8 artikel.

IV. KESIMPULAN

Penggunaan metode Naïve bayes, Decision Tree, Neural Network, Random Forest, dan Support Vector Machine sudah banyak digunakan dalam berbagai macam studi. Pada makalah ini telah mengkategorikan metode Naïve bayes, Decision Tree, Neural Network, Random Forest, dan Support Vector Machine pada bidang klasifikasi dan prediksi. Setelah melakukan studi literature review pada lima tahun terakhir dengan rentang 2019 sampai 2023 bidang yang lebih banyak diteliti adalah bidang prediksi dengan jumlah artikel sebanyak 122 dari 243, dari

122 artikel tersebut 46 diantaranya menggunakan metode Random Forest dan setelah di jumlah kan semua hasil akhir berupa akurasi dan dibagi dengan 46 didapatkan presentase akurasi mencapai 91.18%, selanjutnya 30 artikel menggunakan algoritma Support Vector Machine dengan rata-rata akurasi 89,14%, selanjutnya ada Decision Tree dengan total 28 artikel dan rata-rata akurasi 88,20%, dan dua metode yang jarang digunakan dalam menyelesaikan topik prediksi adalah Naïve Bayes dengan 10 artikel dan rata-rata akurasi 77,51%, dan yang terakhir Neural Network dengan jumlah 8 artikel dan rata-rata akurasi 93,54%. Sedangkan pada bidang klasifikasi pada lima tahun terakhir dari rentang 2019 sampai 2023, Support Vector Machine merupakan metode yang paling tend digunakan dengan jumlah 33 dari 121 artikel dan presentase akurasi 88.85%, selanjutnya ada Neural Network dengan jumlah 26 artikel, setelahnya ada Naïve Bayes dengan jumlah 27 dan rata-rata akurasi 87,45%, selanjutnya Random Forest dengan jumlah artikel sebanyak 22 dan rata-rata akurasi 92,24% dan metode yang jarang digunakan dalam menyelesaikan topik klasifikasi adalah Decision Tree dengan total 13 artikel dan rata-rata akurasi 91,56%, dari penjabaran diatas bisa dilihat jika metode Random Forest bisa menjadi salah satu pilihan dalam menyelesaikan topik prediksi, dan untuk topik klasifikasi Support Vector Machine lah yang menjadi metode paling tren pada lima tahun terakhir dalam rentang tahun 2019 hingga 2023. Diharapkan buah dari penelitian ini juga bisa menjadi acuan dalam memilih metode dalam meneliti sebuah kasus dan penelitian ini juga bisa dikembangkan lagi dengan membandingkan lebih banyak metode dan memperbanyak jumlah artikel yang digunakan dalam literature riview sehingga bisa didapatkan hasil yang lebih akurat.

REFERENSI

- [1] D. Sayhidin, G. Haris, and C. Juliane, "Implementasi Data Mining Tingkat Kepemimpinan Siswa dengan K- Nearest Neighbor , Decision Tree , dan Naïve Bayes," vol. 7, pp. 199–206, 2023, doi: 10.30865/mib.v7i1.5351.
- [2] S. Saadah and H. Salsabila, "Prediksi Harga Ponsel Menggunakan Metode Random Forest," *J. Komput. Terap.*, vol. 7, no. 1, pp. 24–32, 2021.
- [3] H. D. Wijaya and S. Dwiasnati, "Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat," *J. Inform.*, vol. 7, no. 1, pp. 1–7, 2020, doi: 10.31311/ji.v7i1.6203.
- [4] M. Qasim and S. D. Khan, "Detection and Relative Quantification of Neodymium in Sillai Patti Carbonatite Using Decision Tree Classification of the Hyperspectral Data," 2022.
- [5] X. Yan, "Garment Design Models Combining Bayesian Classifier and Decision Tree Algorithm," vol. 2022, 2022.
- [6] A. P. Wibawa, M. G. A. Purnama, M. F. Akbar, and F. A. Dwiyanto, "Metode-metode Klasifikasi," *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 134, 2018.
- [7] M. Muhathir, M. H. Santoso, and D. A. Larasati, "Wayang Image Classification Using SVM Method and GLCM Feature Extraction," *J. Informatics Telecommun. Eng.*, vol. 4, no. 2, pp. 373–382, 2021, doi: 10.31289/jite.v4i2.4524.
- [8] L. Setiyani, "ANALISIS PREDIKSI KELULUSAN MAHASISWA TEPAT WAKTU MENGGUNAKAN METODE DATA MINING NAÏVE BAYES: SYSTEMATIC REVIEW," vol. 13, no. 1, pp. 35–43, 2020, doi: 10.30998/faktorexacta.v13i1.5548.
- [9] M. Guntur, J. Santony, and Y. Yuhandri, "Prediksi Harga Emas dengan Menggunakan Metode Naïve Bayes dalam Investasi untuk Meminimalisasi Resiko," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 2, no. 1, pp. 354–360, 2018, doi: 10.29207/resti.v2i1.276.
- [10] M. I. Aziz and A. Z. Fanani, "Analisis Metode Ensemble Pada Klasifikasi Penyakit Jantung Berbasis Decision Tree," vol. 7, pp. 1–12, 2023, doi: 10.30865/mib.v7i1.5169.

- [11] P. Handayani, E. Nurlelah, M. Raharjo, and P. M. Ramdani, "Prediksi Penyakit Liver Dengan Menggunakan Metode Decision Tree dan Neural Network," *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, p. 55, 2019, doi: 10.24114/cess.v4i1.11528.
- [12] R. Meenal, P. A. Michael, D. Pamela, and E. Rajasekaran, "Weather prediction using random forest machine learning model," vol. 22, no. 2, pp. 1208–1215, 2021, doi: 10.11591/ijeecs.v22.i2.pp1208-1215.
- [13] H. Tuhuteru, "Analisis Sentimen Masyarakat Terhadap Pembatasan Sosial Berksala Besar Menggunakan Algoritma Support Vector Machine," vol. 4, no. 1, 2020.
- [14] R. S. Wahono, "A Systematic Literature Review of Software Defect Prediction : Research Trends , Datasets , Methods and Frameworks," vol. 1, no. 1, 2015.
- [15] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009, doi: 10.1016/j.infsof.2008.09.009.
- [16] S. E. Bailey, S. S. Godbole, C. D. Knutson, and J. L. Krein, "A Decade of Conway ' s Law : A literature review from 2003-2012," pp. 1–14, 2013, doi: 10.1109/RESER.2013.14.
- [17] W. Beitrag, "der Lendenwirbelsa," vol. 48, pp. 43–48, 2007, doi: 10.1016/j.orthtr.2007.01.004.
- [18] M. Lestandy, A. Abdurrahim, and L. Syafa, "Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent," vol. 5, no. 10, pp. 802–808, 2021.
- [19] V. Nhu, A. Shirzadi, H. Shahabi, and S. K. Singh, "Shallow Landslide Susceptibility Mapping : A Comparison between Logistic Model Tree , Logistic Regression , Naïve Bayes Tree , Artificial Neural Network , and Support Vector Machine Algorithms," no. February, 2020.
- [20] S. Soim, P. N. Sriwijaya, A. Info, D. T. Algorithm, and P. Datasets, "Implementation of Decision Tree Algorithm Machine Learning in Detecting Covid-19 Virus Patients Using Public Datasets," vol. 5, no. 1, pp. 37–43, 2022.
- [21] J. Han, M. Fang, S. Ye, C. Chen, Q. Wan, and X. Qian, "Using Decision Tree to Predict Response Rates of Consumer Satisfaction , Attitude , and Loyalty Surveys," 2019.
- [22] M. Aji, P. Wibowo, M. Bima, A. Fayyadl, Y. Azhar, and Z. Sari, "Classification of Brain Tumors on MRI Images Using Convolutional," vol. 5, no. 158, pp. 2–6, 2022.
- [23] M. A. Amrustian and M. Wibowo, "Implementasi Metode Convolutional Neural Network untuk Klasifikasi Breast Cancer pada Citra Histopatologi," vol. 7, pp. 41–49, 2023, doi: 10.30865/mib.v7i1.5194.
- [24] Z. Zhang, M. Wei, D. Pu, G. He, and G. Wang, "Assessment of Annual Composite Images Obtained by Google Earth Engine for Urban Areas Mapping Using Random Forest," 2021.
- [25] F. Wang, Y. Wang, X. Ji, and Z. Wang, "Effective Macrosomia Prediction Using Random Forest Algorithm," 2022.
- [26] W. Rizka, U. Fadilah, D. Agfiannisa, and Y. Azhar, "Analisis Prediksi Harga Saham PT . Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine," vol. 5, no. 2, 2020.
- [27] H. Fang, X. Fu, Z. Zeng, and K. Zhong, "An Improved Arithmetic Optimization Algorithm and Its Application to Determine the Parameters of Support Vector Machine," 2022.