

Klasifikasi Pengembalian Radar dari Ionosfer Menggunakan SVM, Naïve Bayes dan Random Forest

Yulia Aryani^{1*}, Arie Wahyu Wijayanto²

^{1,2)}Peminatan Sains Data, Program Studi Komputasi Statistik, Politeknik Statistika STIS
Jl. Otto Iskandardinata, DKI Jakarta, Indonesia 13330

*email: 221810672@stis.ac.id

(Naskah masuk: 2 Januari 2021; diterima untuk diterbitkan: 31 Maret 2021)

ABSTRAK – Klasifikasi merupakan salah satu topik utama dalam data mining atau machine learning. Klasifikasi adalah suatu pengelompokan data dimana data yang digunakan tersebut mempunyai kelas label atau target. Klasifikasi digunakan untuk mengambil data dan ditempatkan kedalam kelompok tertentu. Studi tentang ionosfer penting untuk penelitian di berbagai domain, khususnya dalam sistem komunikasi. Dalam penelitian ionosfer, perlu dilakukan klasifikasi radar yang berguna dan tidak berguna dari ionosfer. Pada makalah ini, akan dilakukan klasifikasi terhadap data inosphere yang diambil dari UCI machine learning repository. Klasifikasi dilakukan dengan menggunakan tiga metode klasifikasi, yakni SVM (Support Vector Machine), Naïve Bayes, dan Random Forest. Hasil dari percobaan ini bisa menunjukkan prediksi dari setiap percobaan dengan tingkat akurasi dan prediksi yang berbeda-beda di setiap metode yang digunakan. Hasil akurasi, presisi, dan recall terbaik didapatkan pada metode Random Forest dengan rasio data latih dan data uji sebesar 85% didapat akurasi dari data uji sebesar 90,57% dengan presisi sebesar 94,12%.

Kata Kunci – Ionosfer; Klasifikasi; SVM; Naïve Bayes; Random Forest.

Classification of Radar Returns from the Ionosphere Using SVM, Naïve Bayes and Random Forest

ABSTRACT – Classification is one of the main topics in data mining or machine learning. Classification is a grouping of data where the data used has a label or target class. Classification is used to collect data and place it into certain groups. The study of the ionosphere is important for research in various domains, particularly in communication systems. In ionosphere research, it is necessary to classify useful and useless radars of the ionosphere. In this paper, we will classify the inosphere data taken from the UCI machine learning repository. Classification is done using three classification methods, namely SVM (Support Vector Machine), Naïve Bayes, and Random Forest. The results of this experiment can show predictions from each experiment with different levels of accuracy and prediction in each method used. The results of the best accuracy, precision, and recall were obtained in the Random Forest method with a ratio of training data and test data of 85%, the accuracy of the test data was 90.57% with a precision of 94.12%.

Keywords - Ionosphere; Classification; SVM; Naïve Bayes; Random Forest.

1. PENDAHULUAN

Klasifikasi merupakan salah satu topik utama dalam *data mining* atau *machine learning*. Klasifikasi adalah menganalisis data menggunakan model yang menggambarkan kelas data. Model atau pengklasifikasi dibangun untuk memprediksi label

kategorik [1]. Klasifikasi digunakan untuk mengambil data dan ditempatkan kedalam kelompok tertentu [2].

Penelitian terkait klasifikasi pernah dilakukan oleh Ghosh, *et al.*, [3] mengusulkan sebuah teknik klasifikasi *novel Neuro-fuzzy*, di mana metode yang diusulkan lebih unggul dibandingkan

algoritma RBFNN dan ANFIS. Kasinanthan, *et al.*, [4] klasifikasi pada sembilan dan 24 kelas serangga dari dataset Wang dan Xie menggunakan *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), *K-Nearest Neighbours* (KNN), *Naive bayes* (NB), dan *Convolutional Neural Network* (CNN). Dengan *9-fold cross-validation*, tingkat klasifikasi tertinggi 91,5% dan 90% dicapai menggunakan model CNN. Caxinha, *et al.*, [5] klasifikasi katarak berdasarkan teknik *ultrasound* menggunakan klasifikasi Bayes, *K-Nearest Neighbours* (KNN), *Fishes Linear Discriminant* (FLD), dan *Support Vector Machine* (SVM) dengan hasil SVM menunjukkan kinerja tertinggi 90.62%.

Studi tentang ionosfer penting untuk penelitian di berbagai domain, khususnya dalam sistem komunikasi. Dalam penelitian ionosfer, perlu dilakukan klasifikasi radar yang berguna dan tidak berguna dari ionosfer. Pengembalian radar yang berguna dapat digunakan untuk analisis lebih lanjut dan pengembalian radar yang tidak berguna dapat dibuang. Data yang digunakan dalam penelitian ini merupakan data ionosfer yang diambil dari *UCI Machine Learning Repository*. Data radar dikumpulkan dari susunan bertahap enam belas antena frekuensi tinggi. Targetnya adalah elektron bebas di atmosfer [6].

Radar (*Radio Detection and Ranging*) adalah sistem elektromagnetik untuk mendeteksi, mengukur jangkauan dengan memancarkan gelombang radio, menemukan, mencitrakan dan mengidentifikasi target [7]. Radar beroperasi dengan memancarkan gelombang elektromagnetik dan mendeteksi gema yang kembali dari objek refleksi (target) [8].

Ionosfer adalah lapisan atmosfer yang terletak di atas 80 KM dari pusat bumi, terdiri dari atom terionisasi yang secara langsung menyerap radiasi matahari dan menyebabkan aurora. Ionosfer memungkinkan perambatan jarak jauh dengan memantulkan gelombang radio di ketinggian antara 90 KM dan 350 KM di atas permukaan bumi [9].

Penelitian terdahulu pada pengembalian radar dari ionosfer tentang klasifikasi telah dilakukan oleh Sigilito *et al.*, [10] menggunakan *Neural Network*. Hasil menunjukkan, *Multilayer Feedforward Networks* (MLFN'S) mengungguli *Single-Layer Networks*, dengan akurasi 100% pada data latih dan 98% akurasi untuk data uji dibandingkan dengan *Single-Layer Networks* dengan akurasi 94% dan 92%. Adhikari, *et al.*, [11] menggunakan *Random forest* dan *Neural Network* setelah validasi silang 10 kali, akurasi tertinggi oleh *Random Forest* adalah 94,22% dan akurasi untuk *Single-Layer Neural Network* adalah 99,43%. Owen, [12] klasifikasi pada data ionosfer dengan *Single Nearest Neighbor* mendapat akurasi 92,15 pada set uji, *Tubular Neighbor* model pilihan, mendapat akurasi 94,7% dan nilai *Polymars* akurasi

88,1% pada data uji. Campbell, [13] menggunakan SVM pada dataset ionosfer dengan *soft margin* didapat performa terbaik dengan galat *non-zero training* dari 1% dengan performa umum yang sesuai sebesar 96,0%. Ini melebihi *hard margin* ($C = \sim$) performa 92,0% dan lebih baik dibandingkan dengan performa untuk *K-Nearest Neighbours* (92.1%) dan *Quinlan's C4.5* (94.1%).

Berdasarkan penelitian sebelumnya, tujuan dari penelitian ini adalah melakukan perbandingan metode klasifikasi *data mining* yaitu metode SVM, *Naive Bayes* dan *Random Forest* dengan menemukan akurasi dan presisi yang lebih baik untuk mengklasifikasikan pengembalian radar dari ionosfer.

2. TINJAUAN PUSTAKA

Pengembalian radar digunakan untuk mempelajari Fisika Ionosfer pada lapisan E dan F (ketinggian 100 hingga 500 km). Targetnya, elektron bebas di ionosfer, memiliki ukuran penampang melintang kecil pada orde 10^{-30} m^2 . Kerapatan elektron pada bilangan tipikal akan berada di urutan $10^8/\text{m}^3$, dan total volume bisa mencapai 10^6 m^3 . Selain itu, proses hamburan balik disebut koheren, dan sinyal hamburan balik sebanding dengan kuadrat kerapatan bilangan. Rasio *signal-to-noise* biasanya berada dalam kisaran 10 hingga 20dB, tetapi dapat mencapai 50 dB [11].

Radar beroperasi dengan mentransmisikan pola *multipulse* ke ionosfer. Penerima dihidupkan di antara *pulse*, dan kecepatan target ditentukan dengan mengukur pergeseran fase pengembalian.

Gelombang radar merambat melalui atmosfer dengan kecepatan sekitar 300.000 km per detik (kecepatan cahaya). Jangkauan ke target ditentukan dengan mengukur waktu yang dibutuhkan sinyal radar untuk bergerak ke target dan kembali. Kisaran ke target sama dengan $ct/2$, di mana c adalah kecepatan propagasi energi radar dan T adalah waktu perjalanan pulang-pergi yang diukur oleh radar. Akurasi jangkauan radar *pulse* sederhana bergantung pada lebar *pulse*, semakin pendek *pulse*, semakin baik akurasinya. Namun, *pulse* pendek membutuhkan *bandwidth* yang luas dalam penerima dan pemancar (karena *bandwidth* sama dengan kebalikan dari lebar *pulsa*) [14].

3. METODE BAHAN

Data

Data yang digunakan dalam penelitian ini merupakan data ionosfer yang diambil dari *UCI Machine Learning Repository* yang digunakan pada penelitian sebelumnya [11]. Data tersebut berisikan data radar yang digunakan dalam penelitian yang

dikumpulkan oleh Kelompok Fisika Ruang Angkasa dari Laboratorium Fisika Terapan Johns Hopkins University. System radar berada di Goose Bay, Labrador, terdiri dari array bertahap dari 16 antena fekuensi tinggi, dengan total daya yang ditransmisikan 6,4 Kw dan penguat antena sekitar 30 dBm pada rentang frekuensi 8 hingga 20 MHz.

Data terdiri dari 351 baris dan 35 atribut, dengan 34 atribut merupakan atribut prediksi dengan data kontinu dan atribut ke 35 berisi data kategorik "good" dan "bad" yang mendefenisikan nilai target dapat dilihat pada Gambar 1.



Gambar 1. Perbandingan kategori pengembalian radar

Berdasarkan Gambar 1, pada data ionosfer didapat data dengan kategori "good" berjumlah 225 dan data dengan kategori "bad" berjumlah 126.

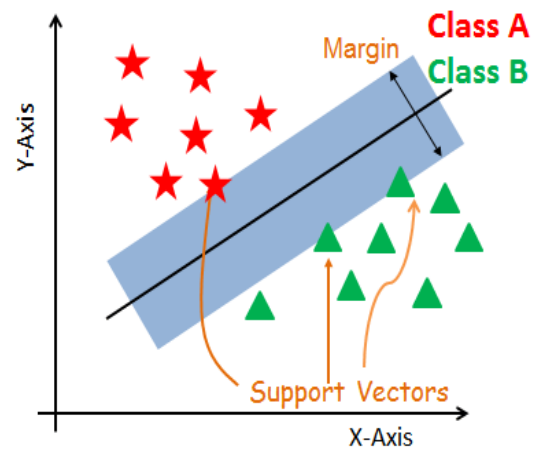
Support Vector Machine (SVM)

SVM adalah metode *supervised machine learning* yang digunakan untuk klasifikasi data. Dalam teknik ini, setiap objek statistik diplotkan sebagai titik dalam ruang berdimensi-n dengan nilai setiap fungsi menjadi biaya koordinat yang dipilih. Selanjutnya dilakukan klasifikasi dengan mendesain *hyperplane* yang membedakan kedua kelas tersebut. Dalam area dua dimensi, *hyperline* adalah garis yang membagi bidang menjadi dua bagian atau kelas. SVM bertujuan untuk mengoptimalkan lebar celah antar kelas [15]. Ide utama dari SVM adalah menemukan *hyperlane* marginal maksimum (MMH) yang paling baik membagi dataset menjadi beberapa kelas ditunjukkan pada Gambar 2 [16].

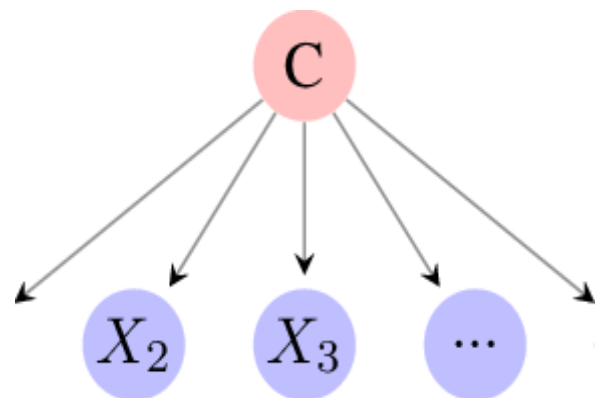
Naïve Bayes

Naïve Bayes (NB) adalah salah satu algoritma data mining yang paling terkenal untuk klasifikasi [17]. *Naïve Bayes Classifier* (NB) adalah pengklasifikasi yang sangat sederhana berdasarkan teorema Bayes. Strukturnya ditunjukkan pada Gambar 3 dimana node C mewakili kelas yang berbeda dan $X_1, X_2, .. X_n$ mewakili komponen atau fitur yang berbeda dari

sebuah sampel. NB mengasumsikan semua *node* fitur di dalam kelas independen dan fitur variabel diasumsikan distribusi *Gaussian* jika besifat kontinu [18].



Gambar 2. Teknik klasifikasi SVM [16]



Gambar 3. Klasifikasi Naive Bayes [18].

Random Forest

Random Forest adalah kombinasi pohon prediktor dimana setiap pohon bergantung pada nilai vektor acak yang dijadikan sampel independen dengan distribusi yang sama untuk semua pohon di hutan. *Random forest* adalah pengklasifikasi yang terdiri dari kumpulan klasifikasi pohon-terstruktur $\{ h(x, \Theta_k), k = 1, .. \}$ di mana $\{ \Theta_k \}$ adalah distribusi vector acak *independent* dan setiap pohon memberikan satu unit suara untuk kelas paling populer pada input x [19].

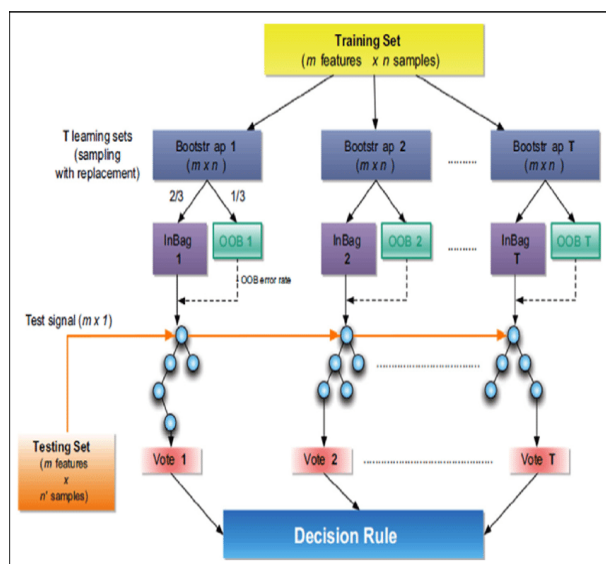
Metode *random forest* merupakan metode klasifikasi yang dilakukan dengan menggunakan struktur pohon dalam jumlah besar daripada satu struktur pohon. Dalam metode ini, sampel dari kumpulan data dipilih dengan metode *Bootstrap* dan pohon klasifikasi dibuat. Langkah-langkah klasifikasi yang dilakukan dengan menggunakan

algoritma random forest ditunjukkan pada Gambar 4 [20].

Langkah Analisis

Langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut.

- 1) Melakukan *preprocessing* data seperti pengecekan *missing value*, mengatasi *outlier*, dan sebagainya.
- 2) Melakukan pemilihan fitur.
- 3) Melakukan validasi dengan *Split Validation*.
- 4) Mengklasifikasikan data menggunakan metode SVM, *Naïve Bayes*, dan *Random Forest*.
- 5) Menginterpretasikan hasil analisis.
- 6) Melakukan perbandingan
- 7) Menarik kesimpulan.



Gambar 4. Flow chart random forest [20]

4. HASIL DAN PEMBAHASAN

Preprocessing Data

Pada data ionosfer tidak didapatkan adanya *missing value*, dan juga tidak terdapat outlier sehingga tidak diperlukannya penanganan *missing value*. Atribut ke 35 merupakan data kategorik sehingga dapat diubah menjadi data faktor untuk dikonversikan ke dalam kode numerik, 1 untuk "good" dan 2 untuk "bad".

Feature Selection

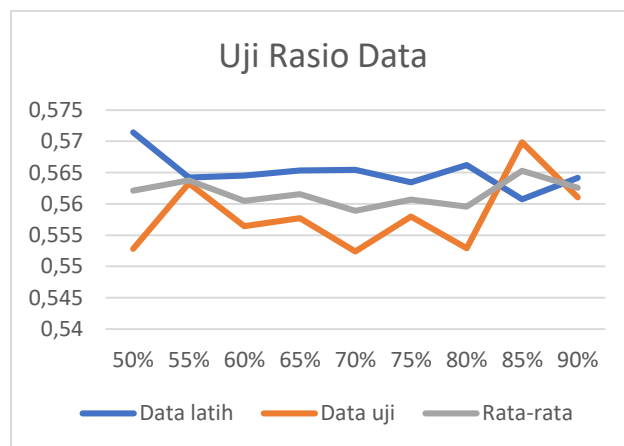
Pada proses ini dilakukan pembuangan variabel kedua (V2) dikarenakan variabel 2 tidak memiliki korelasi dengan variabel manapun sehingga

diperlukan pembuangan variabel agar tidak terjadi kesalahan dalam proses analisis.

Split Validation

Melakukan validasi sederhana dengan membagi dataset secara acak menjadi dua data terpisah data latih dan data uji. Rasio *split* sangat penting untuk menemukan akurasi data *training* dan *testing* yang paling baik, sehingga diperlukan pengujian rasio data.

Pada pengujian pertama digunakan untuk mengetahui perbandingan tingkat akurasi dari dataset yang digunakan. Dataset yang digunakan baik data latih maupun data uji yang berjumlah sebanyak 451 dan dibagi sesuai dengan perbandingan rasio yang telah ditentukan guna untuk melihat hasil akurasi terbaik. Perbandingan rasio yang akan digunakan yaitu rasio 50%, 60%, 70%, 75%, 80% dan 90%. Pengujian ini menggunakan metode SVM dengan tipe SVM *C-Classification*, kernel radial dan *cost* 1 dengan 10 kali pengulangan.



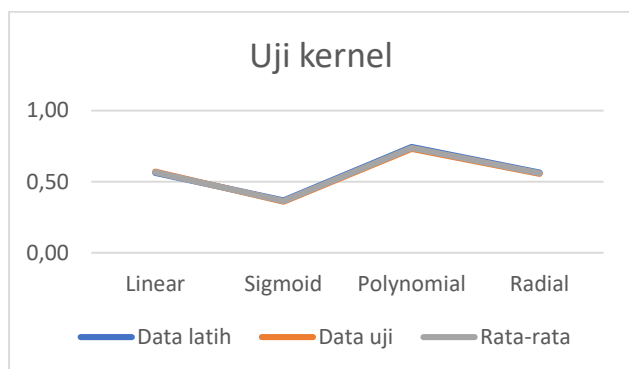
Grafik 1 Perbandingan Rasio Data

Berdasarkan Grafik 1, diperoleh akurasi terbaik terdapat pada rasio data 85% dengan rata-rata akurasi data latih dan data uji 56,5%. Data latih dan data uji yang digunakan dipilih secara acak dari setiap kelas, pada rasio 85% didapat data latih sebanyak 298 dan data uji sebanyak 53.

Klasifikasi dengan metode SVM

Sebelum dilakukan klasifikasi, dilakukan pengujian terhadap jenis kernel yang digunakan, dimana pengujian tersebut dilakukan guna untuk mengetahui jenis kernel yang mana yang memperoleh hasil terbaik. Nilai parameter yang digunakan pada pengujian SVM ini adalah nilai dengan rasio data 85%, tipe SVM *C-Classification*, dan *cost* 1 dan dilakukan pengulangan 10 kali. Berikut adalah hasil rata-rata akurasi dari tiap kernel.

Berdasarkan Grafik 2, diperoleh nilai akurasi yang paling optimal terdapat pada kernel *polynomial* dengan rata-rata nilai akurasi sebesar 73,7%. Hal ini dapat disimpulkan bahwa kernel *polynomial* lebih cocok digunakan pada data *ionosphere* dibandingkan dengan jenis kernel yang lainnya seperti kernel *linear*, *sigmoid*, dan kernel *radial*. Berikut hasil *confusion matrix* dengan metode SVM dengan menggunakan *set.seed* 123 (Tabel 1,2).



Grafik 2 Perbandingan kernel

Tabel 1. Matriks data latih metode SVM

Prediksi	Referensi	
	Good	Bad
Good	113	78
Bad	0	107

Tabel 2. Matriks data uji metode SVM

Prediksi	Referensi	
	Good	Bad
Good	21	13
Bad	0	19

Klasifikasi dengan metode Naïve Bayes

Berikut hasil prediksi dengan metode Naïve Bayes, Pada data latih dengan *laplace* 0, didapat *prior probability* 64,09% untuk klasifikasi *good* yang artinya peluang data kategori *good* untuk prediksi *good* adalah 64,09% dan 35,91% untuk klasifikasi *bad* artinya. Peluang data kategori *bad* untuk diprediksi *bad* ada 35,91%. Berikut *confusion matrix* pada data uji.

Tabel 3. Matriks data uji metode Naive bayes

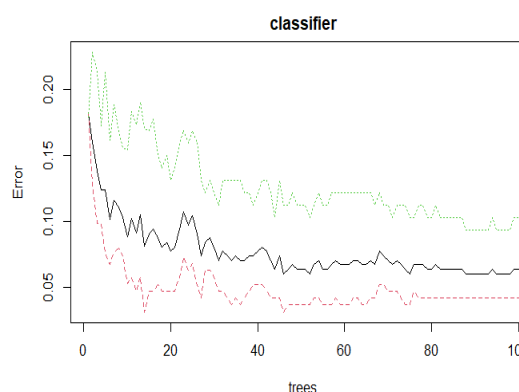
Prediksi	Referensi
----------	-----------

	Good	Bad
Good	23	1
Bad	11	18

Berdasarkan tabel 3 didapat bahwa, dari data test 23 kategori *good* yang di prediksi *good* dan 1 yang diprediksi *bad*, dan juga dari kategori *bad* 11 diprediksi kategori *bad* dan 18 yang diprediksi *good*.

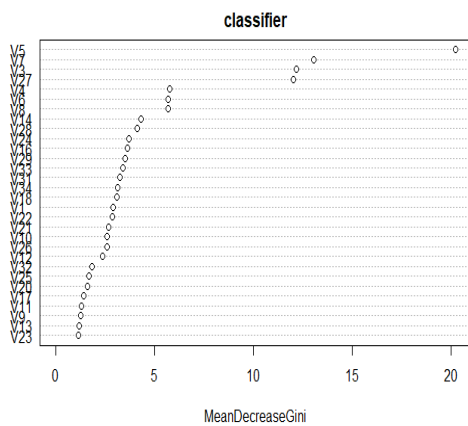
Klasifikasi dengan metode Random Forest

Klasifikasi dilakukan dengan menggunakan *ntree* 100 yang artinya pohon yang akan ditumbuhkan 100 pohon, dan jumlah *variable* yang dipakai setiap percabangan ada 5 dan juga *OOB (Out of Box) estimate* dari *error rate* sebesar 6,35% yang artinya estimasi error yang didapatkan jika data diluar dari data pengujian sebesar 6,35%. Hasilnya berupa 194 data yang benar diprediksi dan 8 data lainnya salah prediksi pada kategori *good* dengan galat 3,96%. Pada kategori *bad* terdapat 99 data yang benar dan 14 data tidak tepat dengan galat 12,39%. Disimpulkan bahwa *random forest* belum mampu 100% memprediksi dengan baik. Berikut diberikan hasil *plot OOB* (Gambar 5) dan *plot MDG* (Gambar 6).



Gambar 1. Plot OOB

Berdasarkan Gambar 5, pada sumbu x merupakan pohon klasifikasi dan pada sumbu y merupakan galat atau tingkat kesalahan. Nilai galat OOB diwarnai dengan warna hitam, ketepatan kategori *good* dengan warna merah dan kategori *bad* dengan warna hijau. Pada *plot* diatas pada pohon ke 60 pelatihan model sudah mulai stabil. Sehingga dapat digunakan untuk klasifikasi secara umum.



Gambar 2. Plot MDG

Berdasarkan Gambar 6, didapat variabel predictor dengan MDG tertinggi adalah V5. Berikut hasil *confusion matrix* data uji dengan metode *Random Forest*.

Tabel 4. Matriks data uji metode random forest

Prediksi	Referensi	
	Good	Bad
Good	32	2
Bad	3	16

Berdasarkan tabel 4 didapat bahwa, dari data test 32 kategori *good* yang di prediksi *good* dan 2 yang diprediksi *bad*, dan juga dari kategori *bad* 16 diprediksi kategori *bad* dan 3 yang diprediksi *good*.

Perbandingan Klasifikasi Metode SVM, Naïve Bayes dan Random Forest.

Tabel 5. Tabel perbandingan prediksi dengan data uji

	SVM	Naïve Bayes	Random Forest
Akurasi	0,7547	0,7736	0,9057
Sensitifitas	1	0,6765	0,9143
spesifisitas	0,5938	0,9474	0,8889
Presisi	0,6176	0,6765	0,9412

Berdasarkan tabel 5, nilai akurasi dari klasifikasi yang dibentuk dengan metode *Random Forest* bernilai 0,9057. Artinya, kemampuan klasifikasi yang dibentuk dengan metode *Random Forest* mengklasifikasikan data pada data uji dengan benar sebesar 90,57%. Sehingga klasifikasi lumayan tepat mengklasifikasikan data pada data latih. Kemudian nilai sensitifitas bernilai 0,9243, artinya proporsi pengembalian radar dari ionosfer memiliki kategori bagus juga diklasifikasikan bagus adalah 92,43%. Begitupula dengan spesifikasi yang bernilai 0,8889, artinya proporsi pengembalian radar dari ionosfer memiliki kategori bad, juga diklasifikasikan *bad* adalah 88,89%. Presisi sebesar 0,9412 artinya

proporsi dari pengembalian radar dari ionosfer yang memiliki kategori *good*, dari seluruh pengembalian yang berkategori *good* adalah 94,12%.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan pada bagian sebelumnya, dapat ditarik kesimpulan bahwa dibandingkan dengan SVM dan *Naive Bayes* metode *Random Forest* merupakan metode yang paling baik digunakan dalam mengklasifikasikan data Ionosfer dari *UCI Machine Learning Repository*. Hal tersebut dikarenakan metode *Random forest* dapat menghasilkan nilai akurasi, dan presisi yang cukup tinggi dan stabil pada data uji, yakni 90,57% pada akurasi dan 94,12% presisi. Berdasarkan penelitian terdahulu ternyata *Neural Network* memiliki akurasi yang lebih tinggi.

Saran yang dapat dilakukan bagi peneliti lain adalah menggunakan metode *Neural Network* untuk data yang mempelajari tentang radar karena menghasilkan nilai akurasi, presisi, dan penarikan yang lebih tinggi. Semakin tinggi ketiga nilai tersebut maka semakin bagus metode dan model yang digunakan.

DAFTAR PUSTAKA

- [1] R. A. Baeza-Yates and B. A. Ribeiro-Neto, *Modern Information Retrieval: The Concept and Technology Behind Search*. 2nd ed, Pearson, 2011.
- [2] Liane Colonna, A Taxonomy and Classification of Data Mining, 16 *SMU Sci. & Tech. L. Rev.* 309, 2013, pp. 342.
- [3] S. Ghosh, S Biswas, D. Sarkar, P. P. Sarkar, 2014, A novel Neuro-fuzzy classification technique for data mining. in *Egyptian Informatics Journal*. vol. 15, issue. 3, pp. 129-147.
- [4] T. Kasinathan, D. Singaraju, S. R. Uyyala, "Insect classification and detection in field crops using modern machine learning techniques", *Journal Information Processing in Agriculture*, Vol. 7, no. 4, pp. 1-12, 2020.
- [5] M. Chaixinha, E.Velte, "Automatic Cataract Classification based on Ultrasound Technique using Machine Learning : A comparative Study". Presented at 2015 *International Congress on Ultrasonics, 2015 ICU Mets*. in *Physics Procedia*. vol. 70. 2015,
- [6] J. G. dy and C. E. Brodley, "Feature Selection for Unsupervised Learning", *Journal of Machine Learning Research*. vol. 5, pp. 845-889, 2004 .
- [7] Hsueh-Jyh Li and Yean-Woei Kiang, "Radar and Inverse Scattering", in *The Electrical Engineering Handbook*, Cambridge, MA, USA:Academic Press, pp. 671-690, 2005.

- [8] R. E. Rinehart, R. E. "Radar for meteorologists", University of North Dakota, Office of the President. 1991.
- [9] Bin-Yi Liu, "HF over the Horizon Radar System Performance Analysis," Naval Postgraduate School Master's thesis, Monterey CA, September 2007.
- [10] V. G. Sigillito, S. P. Wing, "Classification of radar returns from the ionosphere using neural networks" *Johns Hopkins APL Technical Digest*, vol.10, no. 3, pp. 262-266, 1989.
- [11] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," in *International Conference on Electronics and Sustainable Communication Systems (ICESC)*, , pp. 975-978, 2020.
- [12] A. B. Owen, "Tubular neighbors for regression and classification", Stanford University, 1999.
- [13] C. Campbell, and N. Cristianini, "Simple learning algorithms for training support vector machines. Technical report", University of Bristol, 1998.
- [14] M.I. Skolnik. "Radar", [Online]. Britannica.com. <https://www.britannica.com/technology/radar> [Diakses 1 Januari 2020].
- [15] Y. Zhao, and J. Cen, "Data Mining Applications with R", USA : Academic Press, 181-227. 2014.
- [16] A. Navlani. "Support Vector Machines with Scikit-learn", [Online]. Tersedia : datacamp.com. <https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python> [Diakses 1 Januari 2020].
- [17] X. Wu, V. Kumar, "Top 10 algorithms in data mining". *Knowledge and Information Systems*, vol. 14, no. 1, pp. 1-37, 2008.
- [18] Z. Wang, R. M. Hope, et al., "An EEG Workload Classifier for Multiple Subjects,". In Proc. 33rd Annual International Conference of IEE EBMS. Boston. Massachusetts USA, 2011. [Online]. Available : https://www.researchgate.net/publication/221757708_An_EEG_workload_classifier_for_multiple_subjects [Diakses 1 Januari 2020].
- [19] L. Breiman, "Random forest", *Machine Learning* vol. 45, no. 1, pp. 5-32, 2021.
- [20] N. Morizet, N. Godin, J. Tang, E. Maillet, M. Fregonese, and B. Normand, "Classification of acoustic emission signals using wavelets and random forests: application to localized corrosion," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 1026-1037, 2016.