



Analisis Pengaruh Data Scaling Terhadap Performa Algoritme *Machine Learning* untuk Identifikasi Tanaman

Agus Ambarwari¹, Qadhli Jafar Adrian², Yeni Herdiyeni³

¹Program Studi Informatika, Universitas Teknokrat Indonesia, Bandar Lampung

²Program Studi Sistem Informasi, Universitas Teknokrat Indonesia, Bandar Lampung

³Program Studi Ilmu Komputer, Institut Pertanian Bogor, Bogor

¹ambarwariagus@teknokrat.ac.id, ²qadhliadrian@teknokrat.ac.id, ³yeni.herdiyeni@ipb.ac.id

Abstract

Data scaling has an important role in preprocessing data that has an impact on the performance of machine learning algorithms. This study aims to analyze the effect of min-max normalization techniques and standardization (zero-mean normalization) on the performance of machine learning algorithms. The stages carried out in this study included data normalization on the data of leaf venation features. The results of the normalized dataset, then tested to four machine learning algorithms include KNN, Naïve Bayesian, ANN, SVM with RBF kernels and linear kernels. The analysis was carried out on the results of model evaluations using 10-fold cross-validation, and validation using test data. The results obtained show that Naïve Bayesian has the most stable performance against the use of min-max normalization techniques as well as standardization. The KNN algorithm is quite stable compared to SVM and ANN. However, the combination of the min-max normalization technique with SVM that uses the RBF kernel can provide the best performance results. On the other hand, SVM with a linear kernel, the best performance is obtained when applying standardization techniques (zero-mean normalization). While the ANN algorithm, it is necessary to do a number of trials to find out the best data normalization techniques that match the algorithm.

Keywords: min-max normalization, standardization, zero-mean normalization, machine learning algorithms.

Abstrak

Data scaling memiliki peran penting dalam pra-proses data yang berdampak pada kinerja dari algoritme machine learning. Penelitian ini bertujuan menganalisis pengaruh teknik normalisasi min-max dan standarisasi (normalisasi zero-mean) terhadap performa algoritme machine learning. Tahapan yang dilakukan dalam penelitian ini diantaranya pada dataset fitur venasi daun dilakukan normalisasi data. Hasil dari dataset yang telah dinormalisasi, kemudian diujikan ke empat algoritme machine learning antara lain KNN, Naïve Bayesian, ANN, SVM dengan kernel RBF dan kernel linear. Analisis dilakukan pada hasil evaluasi model menggunakan 10-fold cross validation, dan validasi menggunakan data uji. Hasil yang diperoleh menunjukkan bahwa Naïve Bayesian memiliki performa paling stabil terhadap penggunaan teknik normalisasi min-max maupun standarisasi. Untuk algoritme KNN cukup stabil dibandingkan dengan SVM dan ANN. Namun, kombinasi antara teknik normalisasi min-max dengan SVM yang menggunakan kernel RBF dapat memberikan hasil performa terbaik. Disisi lain, SVM dengan kernel linear, performa terbaik diperoleh ketika menerapkan teknik standarisasi (normalisasi zero-mean). Sedangkan algoritme ANN, perlu dilakukan beberapa kali ujicoba untuk mengetahui teknik normalisasi data terbaik yang sesuai dengan algoritme tersebut.

Kata kunci: normalisasi min-max, standarisasi, normalisasi zero-mean, algoritme machine learning.

© 2020 Jurnal RESTI

1. Pendahuluan

Perkembangan *computer vision* dan *machine learning* untuk identifikasi [1] membawa perubahan besar di bidang pertanian [2] dan botani [3], salah satunya dalam identifikasi spesies tanaman secara otomatis [4].

Penggunaan perangkat portabel [5] dan telepon genggam [6] untuk identifikasi tanaman biasanya dengan mengambil gambar daun dari tanaman. Fitur daun yang dapat digunakan untuk identifikasi tanaman diantaranya bentuk [7][8], tekstur [9][10], warna [3],

dan tulang daun atau venasi daun [11]. Dalam satu jenis fitur tersebut memiliki beberapa unit satuan yang berbeda, sehingga perlu dilakukan praproses data.

Praproses data merupakan salah satu tahap dalam *machine learning*. Algoritme *machine learning* belajar dari data, sehingga penting menyiapkan data secara tepat untuk menyelesaikan suatu masalah. Masalahnya, data biasanya tidak tersedia dalam bentuk siap digunakan. Selain itu, algoritme yang berbeda mungkin memerlukan transformasi data yang berbeda. Meski memiliki data yang baik, perlu dipastikan lagi bahwa data tersebut ada dalam skala yang sama. Teknik normalisasi atau *data scaling* memiliki peran penting dalam praproses data [12] dan biasanya digunakan untuk menyamakan skala data agar bernilai antara 0 sampai 1. Penggunaan metode *data scaling* yang tepat dapat mengoptimalkan kinerja dari algoritme *machine learning* [13].

Kombinasi antara teknik normalisasi data dengan algoritme *machine learning* dapat dipastikan memiliki pengaruh terhadap performa yang dihasilkan. Li dan Liu (2011) dalam penelitiannya menyatakan, bahwa normalisasi *min-max* memiliki performa yang bagus dalam hal kecepatan, akurasi, dan kuantitas *support vector* pada SVM. Disisi lain, Tang dan Sutskever (2011) menyatakan pada banyak algoritme *machine learning* langkah standar sebelum pelatihan adalah menghapus rata-rata dari data, yang dikenal dengan *zero-mean* atau standarisasi.

Berkaitan dengan identifikasi tanaman menggunakan fitur daun, berbagai jenis algoritme *machine learning* telah diterapkan. Munisa *et al.* (2015) menggunakan KNN untuk mengidentifikasi tanaman berdasarkan fitur bentuk dan histogram warna daun. Pada penelitiannya menerapkan normalisasi antara 0 sampai 1, dan akurasi tertinggi yang diperoleh sebesar 87.3%. Padao dan Maravillas (2015), menerapkan *Naïve Bayesian* untuk mengklasifikasikan tanaman berdasarkan fitur bentuk dan tekstur daun. Berdasarkan kurva ROC, akurasi yang diperoleh tinggi yaitu 0.981. Singh dan Bhamrah (2015), menerapkan ANN untuk identifikasi tanaman berdasarkan fitur bentuk daun. Akurasi yang diperoleh sebesar 98.8%. Penelitian lain oleh Ambarwari *et al.* (2018) yang mengidentifikasi tanaman berdasarkan tipe venasi daun menggunakan SVM. Pada penelitiannya diterapkan normalisasi *min-max* dan akurasi yang dihasilkan adalah 77.57%.

Penelitian ini bertujuan untuk menganalisis performa algoritme *machine learning* terhadap teknik *data scaling* dalam mengidentifikasi spesies tanaman berdasarkan fitur venasi daun. Selanjutnya, penentuan penggunaan teknik *data scaling* dapat dengan tepat diterapkan pada algoritme *machine learning* yang akan digunakan. Pada penelitian ini, digunakan dataset hasil ekstraksi fitur venasi daun yang masing-masing diterapkan teknik normalisasi *min-max* dan standarisasi

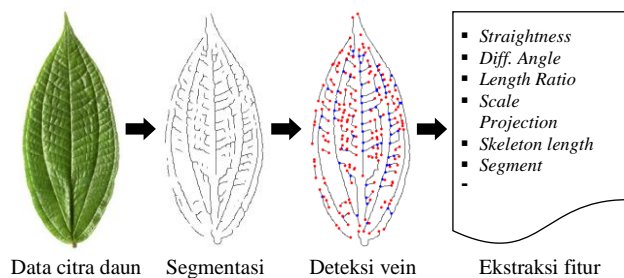
(*zero-mean*) untuk menguji performa 4 algoritme *machine learning*, yaitu *k-nearest neighbor* (KNN), *Naive Bayesian*, *artificial neural networks* (ANN), dan *support vector machine* (SVM). Pemilihan algoritme ini didasarkan pada banyaknya penelitian yang menggunakan algoritme tersebut untuk identifikasi tanaman.

2. Metode Penelitian

Tahapan dalam penelitian ini antara lain: (1) akuisisi data fitur venasi daun; (2) *data scaling* (normalisasi dan standarisasi); (3) klasifikasi; dan (4) analisis hasil. Detail tahapan penelitian dijelaskan pada sub-bab ini.

2.1. Data Fitur Venasi Daun

Data yang digunakan dalam penelitian ini adalah data hasil ekstraksi fitur venasi daun dari dataset Flavia [17], diambil sampel 6 spesies dengan jumlah daun pada setiap spesiesnya 50. Fitur venasi daun terdiri dari 19 fitur, antara lain *mean*, *variance*, *standard deviation (of straightness, different angle, length ratio, scale projection, and length)*, *total skeleton*, *end point*, *branch point*, dan *segment*. Data fitur venasi daun tersebut diperoleh dengan mengikuti tahapan penelitian Ambarwari *et al.* (2018), ilustrasi ditunjukkan pada Gambar 1.



Gambar 1. Tahapan ekstraksi fitur venasi daun

2.2. Data Scaling

Data scaling atau normalisasi data merupakan teknik mengubah nilai numerik dalam dataset ke skala umum, tanpa mendistorsi perbedaan dalam rentang nilai. Normalisasi data akan membantu mempercepat proses pembelajaran pada *machine learning* [13]. Pada data hasil ekstraksi fitur venasi daun selanjutnya dilakukan normalisasi data.

a) Normalisasi *Min-Max*

Normalisasi *min-max* mengubah ukuran data dari rentang asli, sehingga semua nilai berada dalam kisaran 0 dan 1. Persamaannya dapat dilihat pada persamaan (1) [13].

$$v_{norm} = \left(\frac{v_i - v_{min}}{v_{max} - v_{min}} \right) \quad (1)$$

b) Standarisasi (*Zero-Mean*)

Metode normalisasi *Zero-Mean* didasarkan pada *mean* dan standar deviasi. Standarisasi suatu dataset melibatkan perubahan skala distribusi nilai, sehingga nilai rata-rata (*mean*) yang diamati adalah 0 dan standar deviasi adalah 1. Standar deviasi dihitung menggunakan persamaan (2) [13].

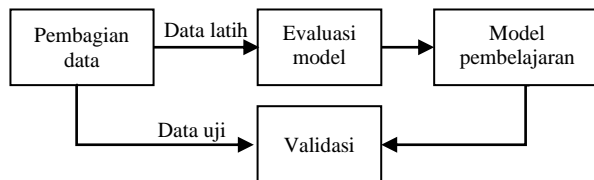
$$x_{std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - x_{mean})^2} \quad (2)$$

x_{mean} adalah rata-rata dari data. Normalisasi dapat dihitung dengan persamaan (3) [13].

$$x'_i = \frac{x_i - x_{mean}}{x_{std}} \quad (3)$$

2.3. Klasifikasi

Beberapa tahapan dalam melakukan pengujian terhadap dataset yang telah dilakukan normalisasi ditunjukkan pada Gambar 2.



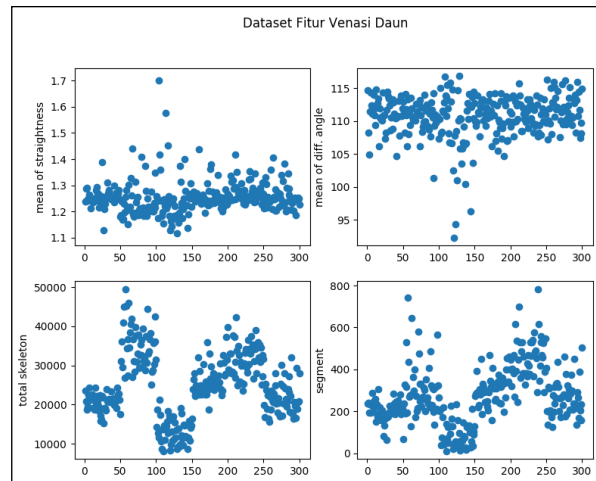
Gambar 2. Tahapan ujicoba *classifier*

Pada dataset hasil normalisasi dilakukan pembagian data, dengan komposisi 75% data latih dan 25% data uji. Dari data latih kemudian dilakukan evaluasi model menggunakan *k-fold cross validation* dengan $k=10$. Evaluasi model dilakukan dengan tujuan mengukur akurasi dari model pada saat pelatihan. Tahap berikutnya dilakukan pelatihan pada data latih, sehingga didapatkan model pembelajaran. Dari model yang telah dibuat, dilakukan validasi menggunakan data uji. Hasil pada evaluasi model dan validasi inilah yang digunakan untuk menganalisis performa dari algoritme *machine learning*.

Dari Gambar 2 pada tahap evaluasi model dan pembuatan model pembelajaran digunakan 4 *classifier*, antara lain KNN, *Naïve Bayesian*, ANN, dan SVM. Ujicoba dilakukan dengan menggunakan bahasa pemrograman Python 3. Adapun algoritme *machine learning* diadopsi dari pustaka *scikit-learn* [18]. Konfigurasi yang dilakukan pada masing-masing algoritme *machine learning* antara lain, pada KNN, nilai *neighbors* yang digunakan adalah $k=3$; *Naïve Bayesian* konfigurasi secara *default*; pada ANN, jumlah *hidden layer*=10, *solver*='lbfgs', *activation*='relu'; sedangkan pada SVM digunakan 2 kernel yaitu linear dan RBF, nilai C dan *gamma* yang digunakan pada kernel RBF masing-masing adalah 1000 dan 0.1.

3. Hasil dan Pembahasan

Pada *machine learning*, setiap dataset tidak selalu memerlukan normalisasi. Normalisasi diperlukan hanya ketika fitur memiliki rentang yang berbeda sebagaimana pada dataset fitur venasi daun. Dataset fitur venasi daun yang digunakan memiliki skala yang berbeda untuk setiap fitur-fiturnya. Mulai dari kisaran 0 sampai puluhan ribu. Visualisasi empat fitur dari dataset fitur venasi daun sebelum dilakukan normalisasi data ditunjukkan pada Gambar 3.



Gambar 3. Visualisasi dataset sebelum dilakukan normalisasi data

Gambar 3 menunjukkan perbedaan skala yang besar antara fitur *mean of straightness* dengan fitur total skeleton. Perbedaan skala yang besar ini dapat meningkatkan komputasi yang dilakukan oleh algoritme *machine learning*. Sehingga untuk mengoptimalkan kinerja dari algoritme *machine learning* perlu dilakukan normalisasi data apabila perbedaan skala pada data cukup besar.

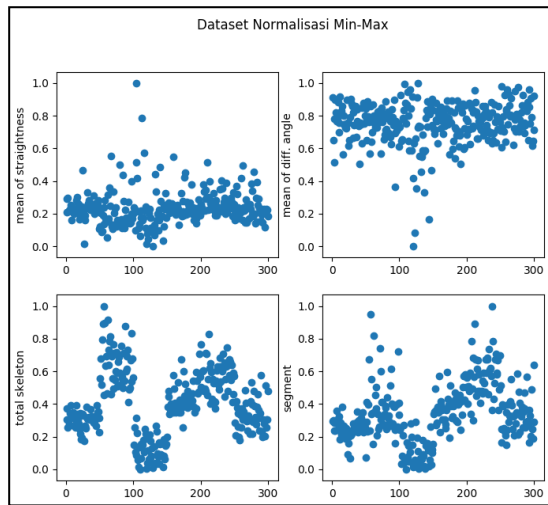
Hasil normalisasi *min-max* pada dataset fitur venasi daun ditunjukkan pada Gambar 4. Normalisasi *min-max* mengubah data kedalam interval 0 sampai dengan 1.

Berbeda dengan normalisasi *min-max*, pada normalisasi *zero-mean* atau standarisasi, perubahan skala dilakukan dengan mengubah nilai rata-rata (*mean*) menjadi 0 dan standar deviasi menjadi 1. Sehingga skala dari setiap fitur masih berbeda (tidak dalam interval yang sama). Visualisasi dataset fitur venasi daun hasil standarisasi ditunjukkan pada Gambar 5.

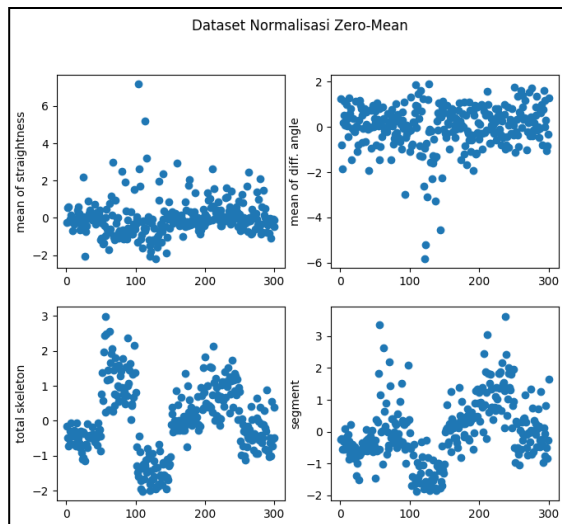
Berdasarkan Gambar 3, Gambar 4, dan Gambar 5, dapat dilihat bahwa pola data tidak ada yang berubah. Perubahan terlihat hanya pada skala setiap fiturnya.

Dataset fitur venasi daun hasil normalisasi (*min-max* dan standarisasi) digunakan untuk menguji performa pada 4 algoritme *machine learning* (KNN, *Naïve Bayesian*, ANN, dan SVM). Pada data latih sebanyak 75% (225 data) dilakukan evaluasi model dengan

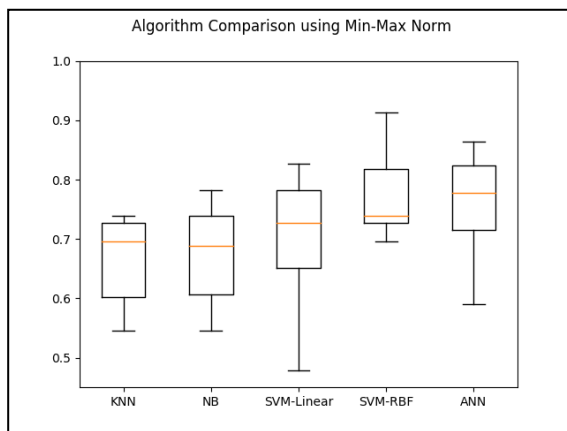
menerapkan *10-fold cross validation*. Hasil evaluasi model pada algoritme *machine learning* yang menggunakan dataset fitur venasi daun hasil normalisasi *min-max* ditunjukkan pada Gambar 6.



Gambar 4. Visualisasi dataset setelah dilakukan normalisasi *min-max*

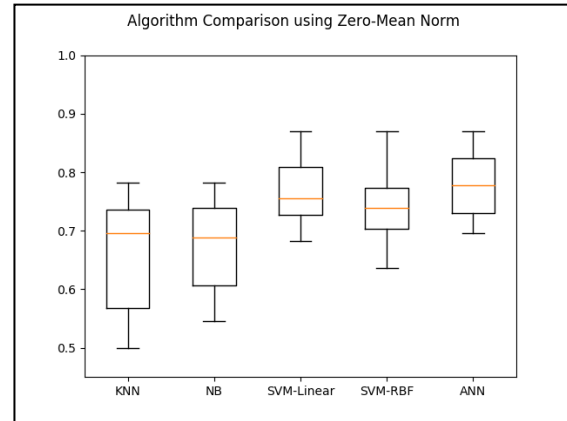


Gambar 5. Visualisasi dataset setelah dilakukan standarisasi



Gambar 6. Evaluasi model menggunakan dataset hasil normalisasi *min-max*

Pada dataset fitur venasi daun hasil normalisasi *min-max* diperoleh rata-rata akurasi tertinggi adalah SVM dengan kernel RBF, yaitu 77.74%. Disusul berikutnya ANN: 76.44%, SVM kernel linear: 70.65%, *Naïve Bayesian*: 67.45%, dan akurasi terendah KNN sebesar 66.62%. Sedangkan hasil evaluasi model dengan menggunakan dataset fitur venasi daun hasil standarisasi (normalisasi *zero-mean*) ditunjukkan pada Gambar 7.



Gambar 7. Evaluasi model menggunakan dataset hasil standarisasi (*zero-mean*)

Pada dataset fitur venasi daun hasil standarisasi (normalisasi *zero-mean*) diperoleh rata-rata akurasi tertinggi sebesar 78.22% dengan algoritme ANN sebagai *classifier*. Berikutnya SVM kernel linear: 76.83%, SVM kernel RBF: 74.60, *Naïve Bayesian*: 67.45%, dan akurasi terendah KNN sebesar 66.16%.

Berdasarkan hasil kedua evaluasi model menggunakan dataset fitur venasi daun hasil normalisasi *min-max* dan standarisasi (*zero-mean*), menunjukkan bahwa KNN dan *Naïve Bayesian* tidak terlalu terpengaruh terhadap normalisasi data. Pengaruh penggunaan normalisasi data terlihat pada SVM kernel linear dan SVM kernel RBF. Pada SVM kernel RBF akurasi tinggi ketika menggunakan dataset normalisasi *min-max*. Sebaliknya pada SVM kernel linear akurasi tertinggi saat dataset dilakukan standarisasi (*zero-mean*). Sedangkan ANN selisih akurasi saat menggunakan kedua dataset hasil normalisasi data tidak terlalu signifikan.

Hasil ujicoba kedua yaitu menggunakan data validasi (data uji) sebanyak 25% (75 data). Data latih 75% (225 data) digunakan untuk pembelajaran pada 4 algoritme *machine learning*, selanjutnya dilakukan validasi menggunakan data uji. Hasil validasi pada dataset fitur venasi daun hasil normalisasi *min-max* ditunjukkan pada Tabel 1.

Berdasarkan Tabel 1, performa terbaik berdasarkan akurasi, rata-rata presisi, dan rata-rata *recall* adalah SVM dengan kernel RBF. Sebagaimana dinyatakan pada penelitian Li dan Liu (2011), penerapan normalisasi *min-max* pada data yang akan diuji

menggunakan SVM kernel RBF memberikan performa yang bagus dalam hal kecepatan dan akurasi [13]. Selain SVM, ANN juga memberikan hasil yang bagus dengan rata-rata seluruhnya lebih dari 81%.

Tabel 1. Perbandingan Algoritme Menggunakan Normalisasi *Min-Max*

<i>Algorithm</i>	<i>Accuracy (%)</i>	<i>Rata-rata Precision (%)</i>	<i>Rata-rata Recall (%)</i>
KNN	77.33	79.00	77.00
Naïve Bayes	70.67	73.00	71.00
SVM Linear	69.34	73.00	69.00
SVM RBF	82.67	84.00	83.00
ANN	81.34	82.00	81.00

Berikutnya hasil validasi pada dataset fitur venasi daun hasil standarisasi (*zero-mean*) ditunjukkan pada Tabel 2.

Tabel 2. Perbandingan Algoritme Menggunakan Normalisasi *Zero-Mean*

<i>Algorithm</i>	<i>Accuracy (%)</i>	<i>Rata-rata Precision (%)</i>	<i>Rata-rata Recall (%)</i>
KNN	76.00	80.00	76.00
Naïve Bayes	70.67	73.00	71.00
SVM Linear	81.34	82.00	81.00
SVM RBF	77.33	78.00	77.00
ANN	76.00	77.00	76.00

Berdasarkan Tabel 2, performa terbaik ditunjukkan pada SVM dengan kernel linear. Berbeda dari Tabel 1, SVM kernel RBF dan ANN mengalami penurunan performa. Dari semua algoritme *machine learning* yang telah diujikan dengan teknik normalisasi data, menunjukkan bahwa KNN dan *Naïve Bayesian* memiliki performa yang stabil. Hal ini dapat dilihat dari hasil evaluasi model dan validasi menggunakan data uji. Kedua algoritme tersebut memberikan hasil akurasi, rata-rata presisi, dan rata-rata *recall* yang hampir sama dari Tabel 1 dan Tabel 2. Bahkan *Naïve Bayesian* memberikan performa paling stabil dari kesemuanya. Sedangkan ANN, performa pada saat evaluasi model dan validasi memberikan hasil yang tidak konsisten. Hal ini dapat dilihat pada Gambar 6 dan Gambar 7, akurasi yang diperoleh memiliki silisih yang tidak signifikan dibandingn pada saat validasi dengan data uji. Performa terbaik diperoleh ketika diterapkan normalisasi *min-max* pada dataset.

4. Kesimpulan

Pengujian dataset fitur venasi daun hasil normalisasi *min-max* dan standarisasi (*zero-mean*) terhadap kelima algoritme *machine learning* diperoleh performa yang bervariasi. Diantara kelima algoritme yang telah diuji, *Naïve Bayesian* memiliki performa yang stabil ketika dataset fitur venasi daun diterapkan normalisasi data, baik normalisasi *min-max* maupun standarisasi (*zero-mean*). Sedangkan KNN, performanya masih cukup stabil dibandingkan SVM dan ANN.

Pada SVM performa terbaik diperoleh ketika mengkombinasikan kernel RBF dengan teknik normalisasi *min-max*. Disisi lain, SVM yang

menerapkan kernel linear, performa terbaik diperoleh ketika dataset dinormalisasi menggunakan teknik standarisasi (*zero-mean*). Sedangkan ANN, perlu dilakukan beberapa kali ujicoba untuk mengetahui teknik normalisasi data terbaik yang sesuai dengan algoritme tersebut.

Adapun saran dari penelitian ini, penggunaan teknik normalisasi memang memiliki pengaruh yang besar terhadap algoritme *machine learning*. Sebaiknya tetap lakukan pengujian dengan sampel data yang bervariasi, karena jumlah data juga memiliki kontribusi dalam menghasilkan performa terbaik algoritme *machine learning*.

Ucapan Terimakasih

Ucapan terima kasih penulis kepada Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia yang telah memberikan dana hibah penelitian melalui Program Penelitian dan Pengabdian Kepada Masyarakat tahun 2018.

Daftar Rujukan

- [1] E. Mata-Montero and J. Carranza-Rojas, "Automated Plant Species Identification: Challenges and Opportunities," *IFIP World Inf. Technol. Forum*, vol. 481, pp. 26–36, 2016.
- [2] E. Hamuda, M. Glavin, and E. Jones, "A survey of image processing techniques for plant extraction and segmentation in the field," *Comput. Electron. Agric.*, vol. 125, pp. 184–199, 2016.
- [3] B. VijayaLakshmi and V. Mohan, "Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features," *Comput. Electron. Agric.*, vol. 125, pp. 99–112, 2016.
- [4] A. Aakif and M. F. Khan, "Automatic classification of plants based on their leaves," *Biosyst. Eng.*, vol. 139, pp. 66–75, 2015.
- [5] J. D. S. Selda, R. M. R. Ellera, L. C. Cajayon, and N. B. Linsangan, "Plant Identification by Image Processing of Leaf Veins," in *ICISPC 2017 Proceedings of the International Conference on Imaging, Signal Processing and Communication*, 2017, pp. 40–44.
- [6] T.-L. Le, N.-D. Duong, V.-T. Nguyen, H. Vu, V.-N. Hoang, and T. T.-N. Nguyen, "Complex Background Leaf-based Plant Identification Method Based on Interactive Segmentation and Kernel Descriptor," in *Proceedings of the 2nd International Workshop on Environmental Multimedia Retrieval*, 2015, pp. 3–8.
- [7] S. Singh and M. S. Bhamrah, "Leaf Identification Using Feature Extraction and Neural Network," *IOSR J. Electron. Commun. Eng.*, vol. 10, no. 5, pp. 134–140, 2015.
- [8] A. Bakhshipour and A. Jafari, "Evaluation of support vector machine and artificial neural networks in weed detection using shape features," *Comput. Electron. Agric.*, vol. 145, pp. 153–160, 2018.
- [9] D. Tomar and S. Agarwal, "Leaf Recognition for Plant Classification Using Direct Acyclic Graph Based Multi-Class Least Squares Twin Support Vector Machine," *Int. J. Image Graph.*, vol. 16, no. 03, p. 1650012, 2016.
- [10] F. R. F. Padoa and E. A. Maravillas, "Using Naïve Bayesian method for plant leaf classification based on shape and texture features," in *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and*

- Control, Environment and Management (HNICEM)*, 2015, no. 12, pp. 117–122, December.
- [11] A. Ambarwari, Y. Herdiyeni, and I. Hermadi, "Identification of Venation Type Based on Venation Density using Digital Image Processing," *J. Teknoinfo*, vol. 12, no. 2, pp. 87–92, 2018.
- [12] A. B. A. Graf and S. Borer, "Normalization in support vector machines," in *Radig B., Florczyk S. (eds) Pattern Recognition*, 2001, pp. 277–282.
- [13] W. Li and Z. Liu, "A method of SVM with normalization in intrusion detection," *Procedia Environ. Sci.*, vol. 11, pp. 256–262, 2011.
- [14] Y. Tang and I. Sutskever, "Data normalization in the learning of restricted Boltzmann machines," in *Department of Computer Science, University of Toronto, Technical Report UTML-TR-11-2*, 2011.
- [15] T. Munisami, M. Ramsurn, S. Kishnah, and S. Pudaruth, "Plant Leaf Recognition Using Shape Features and Colour Histogram with K-nearest Neighbour Classifiers," *Procedia Comput. Sci.*, vol. 58, pp. 740–747, 2015.
- [16] A. Ambarwari, Y. Herdiyeni, and I. Hermadi, "Biometric Analysis of Leaf Venation Density Based on Digital Image," *TELKOMNIKA (Telecommunication Comput. Electron. Control)*, vol. 16, no. 4, p. 1735, 2018.
- [17] Z. Wang, X. Sun, Y. Zhang, Z. Ying, and Y. Ma, "Leaf recognition based on PCNN," *Neural Comput. Appl.*, vol. 27, no. 4, pp. 899–908, 2016.
- [18] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python Fabian," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.