

BAB VI PENUTUP

A. Kesimpulan

Berdasarkan hasil pengujian menggunakan dataset Indeks Standar Pencemaran Udara kota Yogyakarta tahun 2020-2021, menunjukkan bahwa model yang terbaik secara keseluruhan metrik evaluasi adalah *K-Nearest Neighbors* dengan rerata *precision*, *recall*, dan *f-score* sebesar 0,967 0,987, dan 0,977 untuk kelas moderate. Rerata *precision*, *recall*, dan *f-score* yang dihasilkan pada kelas *good* sebesar 0,994, 0,989, dan 0,989 secara berurutan. Rerata *precision*, *recall*, dan *f-score* yang dihasilkan dari kelas *unhealthy* sebesar 1,000, 1,000, dan 1,000 secara berurutan. Rerata *accuracy* keseluruhan dari hasil prediksi model *K-Nearest Neighbors* bernilai sebesar 0,984.

Hasil pengujian yang paling optimal terletak pada model *K-Nearest Neighbors* dengan setelan parameter *nearest neighbors* sebesar 2 dan pembobotan *distanceWeighting* 1 per *distance*. Kelas *moderate* memiliki *precision*, *recall*, dan *f-score* yang bernilai sebesar 0,979, 0,985, dan 0,982 pada kelas *moderate*. Kelas *good* memiliki *precision*, *recall*, dan *f-score* bernilai 0,993, 0,989, dan 0,991. Pada kelas *unhealthy* rerata *precision*, *recall*, dan *f-score* yang dihasilkan bernilai sebesar 1,000 serta *accuracy* yang diperoleh bernilai sebesar 0,988.

B. Saran

Berdasarkan pengujian model pembelajaran mesin yang telah dilaksanakan, penulis menyadari bahwa laporan tugas akhir ini masih jauh dari kata sempurna. Oleh karena itu, saran yang dapat ditampung dalam penelitian ini antara lain adalah:

1. Mengoptimalkan setelan *hyperparameter* lain yang tersedia pada model *Random Forest*, seperti *numFeatures*, *bagSizePercent*, dan lain-lain.
2. Mengoptimalkan setelan *hyperparameter* lain yang tersedia pada model *K-Nearest Neighbors*, terutama memanfaatkan opsi lain *distanceFunction* pada algoritma *nearestSearchAlgorithm* seperti *ManhattanDistance*, *MinkowskyDistance*, *FilteredDistance*, dan *ChebysevDistance*.
3. Mengoptimalkan setelan *hyperparameter K-Nearest Neighbors*, terutama untuk *nearestSearchAlgorithm* dengan memanfaatkan opsi selain *LinearNNSearch*, yaitu *KDTree*, *FilteredNeighborSearch*, *BallTree*, dan *CoverTree*.





DAFTAR PUSTAKA

- [1] A. Fikri, I. A. Darmawan, dan M. Fatkhurrohman, “Rancang Bangun Monitoring Kadar Polusi Udara di Lingkungan Kampus FKIP Menggunakan Sistem IoT,” *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, vol. 6, no. 5, hlm. 2931–2935, Mei 2023, doi: 10.54371/jiip.v6i5.1716.
- [2] T. Kwanda, “Pembangunan Permukiman yang Berkelanjutan untuk Mengurangi Polusi Udara,” Des 2015.
- [3] M. Negara Lingkungan Hidup, “Keputusan Menteri Negara Lingkungan Hidup No. 45 Tahun 1997 Tentang : Indeks Standar Pencemar Udara”.
- [4] A. MM, “Air Quality in Yogyakarta, Indonesia (2021).” Diakses: 14 Desember 2023. [Daring]. Tersedia pada: <https://www.kaggle.com/datasets/adhang/air-quality-in-yogyakarta-indonesia-2021>
- [5] “Air Quality in Yogyakarta, Indonesia (2020).” Diakses: 19 Desember 2023. [Daring]. Tersedia pada: <https://www.kaggle.com/datasets/adhang/air-quality-in-yogyakarta-indonesia-2020>
- [6] A. Aziiz, H. Kirono, I. Asror, Y. Firdaus, dan A. Wibowo, “Klasifikasi Tingkat Kualitas Udara DKI Jakarta Menggunakan Algoritma Naive Bayes,” *EProceedings*, vol. 9, no. 3, hlm. 1962–1969, 2022.
- [7] S. C. H. Nasution, F. Rakhmawati, dan R. S. Lubis, “KLASIFIKASI TINGKAT PENCEMARAN UDARA PADA SEKTOR INDUSTRI DENGAN METODE RANDOM FOREST,” *CSRID (Computer Science Research and Its Development Journal)*, vol. 13, no. 3a, hlm. 53–62, 2021.
- [8] A. Nugroho, I. Asror, dan Y. F. A. Wibowo, “Klasifikasi Tingkat Kualitas Udara DKI Jakarta Berdasarkan Open Government Data Menggunakan Algoritma Random Forest,” *eProceedings of Engineering*, vol. 10, no. 2, 2023.
- [9] A. Amalia, A. Zaidiah, dan I. N. Isnainiyah, “Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 2, hlm. 496–507, 2022.
- [10] Y. Yuliska dan K. U. Syaliman, “Peningkatan Akurasi K-Nearest Neighbor Pada Data Index Standar Pencemaran Udara Kota Pekanbaru,” *IT Journal Research and Development*, vol. 5, no. 1, hlm. 11–18, Jul 2020, doi: 10.25299/itjrd.2020.vol5(1).4680.

- [11] S. Fardiaz, "Mikrobiologi Pangan I. Jakarta: PT. Gramedia Pustaka Utama," *Google Scholar*, 1992.
- [12] M. Soedomo, "Pencemaran udara," *Bandung: ITB*, 2001.
- [13] D. Vallero, *Fundamentals of Air Pollution - Fifth Edition*. 2014. doi: 10.1016/B978-0-12-401733-7.01001-X.
- [14] N. Aini, R. Ruktiari, M. R. Pratama, dan A. F. Buana, "Sistem Prediksi Tingkat Pencemaran Polusi Udara dengan Algoritma Naive Bayes di Kota Makassar Air Pollution Level Prediction System Using Naive Bayes Algorithm in Makassar," *Pros. Semin. Nas. Komun. dan Inform.*, vol. 3, hlm. 83–90, 2019.
- [15] H. Aljuaid dan N. Alwabel, "Air pollution prediction using machine learning algorithms," *Int J Eng Adv Technol*, vol. 8, no. 6 Special Issue 3, hlm. 160–164, 2019.
- [16] A. B. Sianipar dan F. Yuliani, "Optimalisasi Fungsi Papan Indeks Standar Pencemar Udara (Ispu) oleh Dinas Lingkungan Hidup dan Kebersihan Kota Pekanbaru," *Jurnal Online Mahasiswa Fakultas Ilmu Sosial dan Ilmu Politik Universitas Riau*, vol. 4, no. 2, hlm. 1–15, Des 2017.
- [17] A. Kurniawan, "Pengukuran parameter kualitas udara (CO, NO₂, SO₂, O₃ dan PM₁₀) di Bukit Kototabang berbasis ISPU," *Jurnal Teknosains*, vol. 7, no. 1, hlm. 1–13, 2018.
- [18] A. Hermawan, M. Hananto, dan D. Lasut, "Peningkatan Indeks Standar Pencemaran Udara (Ispu) Dan Kejadian Gangguan Saluran Pernapasan Di Kota Pekanbaru," *Indonesian Journal of Health Ecology*, vol. 15, no. 2, hlm. 76–86, 2016.
- [19] E. Alpaydin, *Machine Learning*. The MIT Press, 2021. doi: 10.7551/mitpress/13811.001.0001.
- [20] P. D. Kusuma, *Machine Learning Teori, Program, dan Studi Kasus*. Deepublish, 2020.
- [21] D. Hand, H. Mannila, dan P. Smyth, *Principles of Data Mining*. 2001.
- [22] J. Han, M. Kamber, dan J. Pei, *Data Mining: Concepts and Techniques*. 2012. doi: 10.1016/C2009-0-61819-5.
- [23] W. Supriyanti, K. Kusrini, dan A. Amborowati, "Perbandingan Kinerja Algoritma C4. 5 dan Naive Bayes Untuk Ketepatan Pemilihan Konsentrasi Mahasiswa," *Jurnal Informa: Jurnal Penelitian Dan Pengabdian Masyarakat*, vol. 1, no. 3, hlm. 61–67, 2016.

- [24] S. Moein, *Electrocardiogram signal classification and machine learning: Emerging research and opportunities*. 2018. doi: 10.4018/9781522555803.
- [25] A. I. Sciences, "Machine learning: Random forest with python from scratch© 1st edition: 9781803236803, 9781803236803," *VitalSource*. Packt Publishing. [Daring]. Tersedia pada:
<https://www.vitalsource.com/products/machine-learning-random-forest-with-python-from-ai-sciences-v9781803236803>
- [26] M. Keith, "Random Forest," *Machine Learning with Regression in Python*, 2020, [Daring]. Tersedia pada:
<https://api.semanticscholar.org/CorpusID:19352336>
- [27] L. Breiman, "Random Forest," *Mach Learn*, vol. 45, no. 1, hlm. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [28] S. Zhang, "Challenges in KNN Classification," *IEEE Trans Knowl Data Eng*, vol. 34, no. 10, hlm. 4663–4675, 2022, doi: 10.1109/TKDE.2021.3049250.
- [29] S. Wiyono dan T. Abidin, "Implementation Of K-Nearest Neighbour (KNN) Algorithm to Predict Student's Performance," *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, vol. 9, no. 2, hlm. 873–878, 2018.
- [30] S. Zhang, X. Li, M. Zong, X. Zhu, dan D. Cheng, "Learning k for KNN Classification," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 3, Jan 2017, doi: 10.1145/2990508.
- [31] C. Sammut dan G. Webb, Ed., *Encyclopedia of Machine Learning*. Springer, 2011.
- [32] S. Ruuska, W. Hämmäläinen, S. Kajava, M. Mughal, P. Matilainen, dan J. Mononen, "Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle," *Behavioural Processes*, vol. 148, hlm. 56–62, 2018, doi: <https://doi.org/10.1016/j.beproc.2018.01.004>.
- [33] I. J. Fadillah dan C. D. Puspita, "PEMANFAATAN METODE WEIGHTED K-NEAREST NEIGHBOR IMPUTATION (WEIGHTED KNNI) UNTUK MENGATASI MISSING DATA," *Seminar Nasional Official Statistics*, vol. 2020, no. 1, hlm. 511–518, Jan 2021, doi: 10.34123/semnasoffstat.v2020i1.409.