

**Author Contributions:** Conceptualization, S.M., F.F. and A.; methodology, S.M., F.F., A. and A.H.A.; software, S.M., F.F. and A.; formal analysis, S.M., F.F., A. and A.H.A.; investigation, S.M. and F.F.; validation, S.M. and A.; writing—original draft preparation, S.M., F.F. and A.H.A.; writing—review and editing, S.M., F.F., A. and A.H.A.; project administration, S.M. and A.; funding acquisition, S.M. and A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are all openly available. The dataset for water quality is found at <https://data.jakarta.go.id/dataset/data-hasil-pengujian-kualitas-air-sumur-provinsi-dki-jakarta-tahun-2017>, accessed on 12 April 2023.

**Acknowledgments:** We thank Telkom University for help in the project administration.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Iyer, S.; Kaushik, S.; Chahal, P. Water Quality Prediction Using Machine Learning. *MR Int. J. Eng. Technol.* **2023**, *10*, 59–62. [[CrossRef](#)]
2. Sen, S.; Maiti, S.; Manna, S.; Roy, B.; Ghosh, A. Smart Prediction of Water Quality System for Aquaculture Using Machine Learning Algorithms. *J. Curr. Trends Comput. Sci. Res.* **2023**, *2*, 212–225. [[CrossRef](#)]
3. Xin, L.; Mou, T. Research on the application of multimodal-based machine learning algorithms to water quality classification. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 95557903. [[CrossRef](#)]
4. Patel, J.; Amipara, C.; Ahanger, T.A.; Ladhva, K.; Gupta, R.K.; Alsaab, H.O.; Althobaiti, Y.S.; Ratna, R. A machine learning-based water potability prediction model by using synthetic minority oversampling technique and explainable AI. *Comput. Intell. Neurosci.* **2022**, *2022*, 9283293. [[CrossRef](#)] [[PubMed](#)]
5. Ahmed, U.; Mumtaz, R.; Anwar, H.; Shah, A.A.; Irfan, R.; García-Nieto, J. Efficient Water Quality Prediction Using Supervised Machine Learning. *Water* **2019**, *11*, 2210. [[CrossRef](#)]
6. Wong, W.Y.; Al-Ani, A.; Hasikin, K.; Khairuddin, A.S.M.; Razak, S.A.; Hizaddin, H.F.; Mokhtar, M.I.; Azizan, M.M. Water Quality Index Using Modified Random Forest Technique: Assessing Novel Input Features. *CMES-Comput. Model. Eng. Sci.* **2022**, *132*, 1011–1038.
7. Lemaitre, G.; Nogueira, F.; Aridas, C.K. Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *J. Mach. Learn. Res.* **2017**, *18*, 1–5.
8. Camacho-Navarro, J.; Ruiz, M.; Villamizar, R.; Mujica, L.E.; Moreno-Beltrán, G. Ensemble Learning as Approach for Pipeline Condition Assessment. *J. Phys.* **2017**, *842*, 012019. [[CrossRef](#)]
9. Hordri, N.F.; Yuhaniz, S.S.; Azmi, N.F.M.; Shamsuddin, S.M. Handling Class Imbalance in Credit Card Fraud Using Resampling Methods. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*, 390–396. [[CrossRef](#)]
10. Ajayi, O.; Adetunmbi, A.; Olowookere, T.; Sodiya, S. Performance evaluation of ensemble learning algorithms and classical machine learning algorithms for phishing detection. In Proceedings of the 2002 Smart, Secure and Sustainable Nation, Abuja, Nigeria, 8 November 2022.
11. Prinzie, A.; Van Den Poel, D. Random Forests for Multiclass Classification: Random MultiNomial Logit. *Expert Syst. Appl.* **2008**, *34*, 1721–1732. [[CrossRef](#)]
12. Hancock, J.; Khoshgoftaar, T.M. CatBoost for Big Data: An Interdisciplinary Review. *J. Big Data* **2020**, *7*, 94. [[CrossRef](#)] [[PubMed](#)]
13. Ahn, J.M.; Kim, J.; Kim, K. Ensemble Machine Learning of Gradient Boosting (XGBOOST, LightGBM, CatBoost) and Attention-Based CNN-LSTM for Harmful Algal Blooms Forecasting. *Toxins* **2023**, *15*, 608. [[CrossRef](#)] [[PubMed](#)]
14. Torkey, M.; Bakhiet, A.; Bakrey, M.; Ismail, A.A.; Seddawy, A.I. Recognizing safe drinking water and predicting water quality index using Machine Learning Framework. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 23–33. [[CrossRef](#)]
15. Krtolica, I.; Savić, D.; Bajić, B.; Radulović, S. Machine learning for water quality assessment based on macrophyte presence. *Sustainability* **2022**, *15*, 522. [[CrossRef](#)]
16. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. Performance analysis of the water quality index model for predicting water state using machine learning techniques. *Process Saf. Environ. Prot.* **2023**, *169*, 808–828. [[CrossRef](#)]
17. Kaddoura, S. Evaluation of Machine Learning Algorithm on Drinking Water Quality for Better Sustainability. *Sustainability* **2022**, *14*, 11478. [[CrossRef](#)]
18. Zhu, M.; Wang, J.; Yang, X.; Zhang, Y.; Zhang, L.; Ren, H.; Wu, B.; Ye, L. A review of the application of machine learning in water quality evaluation. *Eco-Environ. Health* **2022**, *1*, 107–116. [[CrossRef](#)]
19. Hassan, N.; Woo, C.S. Machine learning application in water quality using satellite data. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *842*, 012018. [[CrossRef](#)]

20. Bhattacharya, S.; Mishra, S. Applications of Machine Learning for Facies and Fracture Prediction Using Bayesian Network Theory and Random Forest: Case Studies from the Appalachian Basin, USA. *J. Pet. Sci. Eng.* **2018**, *170*, 1005–1017. [CrossRef]
21. Awad, I.; Ridwan, R.L.; Muhammed, M.M.; Abdulaziz, R.O.; Saheed, G.A. Comparison of the CatBoost Classifier with Other Machine Learning Methods. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 738–748. [CrossRef]
22. Data Hasil Pengujian Kualitas Air Sumur Provinsi DKI Jakarta Tahun 2017. Available online: <https://data.jakarta.go.id/dataset/data-hasil-pengujian-kualitas-air-sumur-provinsi-dki-jakarta-tahun-2017> (accessed on 12 April 2023).
23. Google Map. Available online: <https://www.google.com/maps/place/Jakarta,+Daerah+Khusus+Ibukota+Jakarta/@-6.2181616,106.7601824,11.65z/data=!4m6!3m5!1s0x2e69f3e945e34b9d:0x5371bf0fdad786a2!8m2!3d-6.1944491!4d106.8229198!16zL20vMDQ0cnY?hl=id&entry=ttu> (accessed on 4 December 2023).
24. Mairizki, F.; Angga, R.P.; Putra, A.Y. Assessment of groundwater quality for drinking purpose in an industrial area, Dumai City, Riau, Indonesia. *J. Geosci. Eng. Environ. Technol.* **2020**, *5*, 234–238. [CrossRef]
25. Ujianti, R.M.; Anggoro, S.; Bambang, A.N.; Purwanti, F. Water quality of the Garang River, Semarang, Central Java, Indonesia based on the Government Regulation Standard. *J. Phys. Conf. Ser.* **2018**, *1025*, 012037. [CrossRef]
26. Liu, A.Y.; Ghosh, J.; Martin, C.E. Generative Oversampling for Mining Imbalanced Datasets. *Int. Conf. Data Min.* **2007**, *7*, 66–72.
27. Batista, G.E.; Prati, R.C.; Monard, M.C. A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data. *SIGKDD Explor.* **2004**, *6*, 20–29. [CrossRef]
28. Wongvorachan, T.; He, S.; Bulut, O. A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining. *Information* **2023**, *14*, 54. [CrossRef]
29. Menardi, G.; Torelli, N. Training and Assessing Classification Rules with Imbalanced Data. *Data Min. Knowl. Discov.* **2012**, *28*, 92–122. [CrossRef]
30. Bitto, A.K.; Bijoy, M.H.I.; Yesmin, S.; Mahmud, I.; Mia, M.J.; Biplob, K.B.B. Tumor-Net: Convolutional Neural Network Modeling for Classifying Brain Tumors from MRI Images. *Int. J. Adv. Intell. Inform.* **2023**, *9*, 148. [CrossRef]
31. Moradi, S.; Omar, A.; Zhou, Z.; Agostino, A.; Gandomkar, Z.; Bustamante, H.; Power, K.; Henderson, R.; Leslie, G. Forecasting and optimizing dual media filter performance via machine learning. *Water Res.* **2023**, *235*, 119874. [CrossRef]
32. React: The Library for Web and Native User Interfaces. Available online: <https://react.dev> (accessed on 10 June 2023).
33. FastAPI: FastAPI Framework, High Performance, Easy to Learn, Fast to Code, Ready for Production. Available online: <https://fastapi.tiangolo.com> (accessed on 10 June 2023).
34. Python. Available online: <https://www.python.org> (accessed on 10 June 2023).
35. MySQL: The World’s Most Popular Open Source Database. Available online: <https://www.mysql.com> (accessed on 12 June 2023).
36. MQTT: The Standard for IoT Messaging. Available online: <https://mqtt.org> (accessed on 12 June 2023).
37. React Native: Learn Once, Write Anywhere. Available online: <https://reactnative.dev> (accessed on 10 June 2023).
38. Ali, J.; Khan, R.; Ahmad, N.; Maqsood, I. Random Forests and Decision Trees. *Int. J. Comput. Sci. Issues* **2012**, *9*, 272–278.
39. Kulkarni, V.; Sinha, P. Random forest classifiers: A survey and future research directions. *Int. J. Adv. Comput.* **2013**, *9*, 1144–1153.
40. Wen, X.; Xu, P.; Fang, J.; Zhao, Q.; Gu, Z.; Zhang, Q. Multi-Dimensional Data-Based Medium- and Long-Term Power-Load Forecasting Using Double-Layer CatBoost. *Energy Rep.* **2022**, *8*, 8511–8522.
41. Xu, X.; Lai, T.; Jahan, S.; Farid, F.; Bello, A. A machine learning predictive model to detect water quality and pollution. *Future Internet* **2022**, *14*, 324. [CrossRef]
42. Tasnim, A.; Saiduzzaman, M.; Rahman, M.A.; Akhter, J.; Rahaman, A.S.M.M. Performance Evaluation of Multiple Classifiers for Predicting Fake News. *J. Comput. Commun.* **2022**, *10*, 1–21. [CrossRef]
43. Akyol, K.; Atila, Ü. Comparing The Effect of Under-Sampling and Over-Sampling on Traditional Machine Learning Algorithms for Epileptic Seizure Detection. *Acad. Platf.-J. Eng. Sci.* **2020**, *8*, 279–285. [CrossRef]
44. Vilorio, A.; Lezama, O.B.P.; Mercado-Caruzo, N. Unbalanced Data Processing Using Oversampling: Machine Learning. *Procedia Comput. Sci.* **2020**, *175*, 108–113. [CrossRef]
45. Matsuki, K.; Kuperman, V.; Van Dyke, J.A. The Random Forests Statistical Technique: An Examination of Its Value for the Study of Reading. *Sci. Stud. Read.* **2016**, *20*, 20–33. [CrossRef]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.