

## REFERENCES

- Abiad, M., Kadry, S., & Ionescu, S. (2018). Preventive Predictive Maintenance of Telecommunication Equipment - A Review. *Proceedings of the 4th International Conference on Applied and Theoretical Computing and Communication Technology, ICATccT*, 2018, 154–159. <https://doi.org/10.1109/ICATCCT44854.2018.9001972>
- Al Masry, Z., Zemouri, R., Remadna, I., Terrissa, S. L., & Zerhouni, N. (2020). Hybrid Architecture of Deep Convolutional Variational Auto-encoder for Remaining useful Life Prediction. *Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference*, 3592–3598. [https://doi.org/10.3850/978-981-14-8593-0\\_4876-CD](https://doi.org/10.3850/978-981-14-8593-0_4876-CD)
- Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1), 20–29. <https://doi.org/10.1145/1007730.1007735>
- Bloch, H. P., & Carroll, J. R. (1990). Preventive maintenance can be more effective than predictive programs. *Oil and Gas Journal; (USA)*, 88:31.
- Branco, P., Torgo, L., & Ribeiro, R. P. (2017). A Survey of Predictive Modeling on Imbalanced Domains. *ACM Computing Surveys*, 49(2), 1–50. <https://doi.org/10.1145/2907070>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/BF00058655>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cahyana, N., Khomsah, S., & Aribowo, A. S. (2019). Improving Imbalanced Dataset Classification Using Oversampling and Gradient Boosting. *2019 5th International Conference on Science in Information Technology (ICSITech)*, 217–222. <https://doi.org/10.1109/ICSITech46713.2019.8987499>
- Chawla, N. V. (2009). Data Mining for Imbalanced Datasets: An Overview. In *Data Mining and Knowledge Discovery Handbook* (pp. 875–886). Springer US. [https://doi.org/10.1007/978-0-387-09823-4\\_45](https://doi.org/10.1007/978-0-387-09823-4_45)
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>

- Cicak, S., & Avcı, U. (2023). Handling Imbalanced Data in Predictive Maintenance: A Resampling-Based Approach. *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1–6. <https://doi.org/10.1109/HORA58378.2023.10156799>
- Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., & Figari, M. (2016). Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 230(1), 136–153. <https://doi.org/10.1177/1475090214540874>
- Cost Avoidance - Proactive Maintenance Consultants*. (n.d.). Retrieved July 4, 2025, from <https://pmcmn.com/cost-avoidance/>
- Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Dai, Q., Liu, J. wei, & Liu, Y. (2022). Multi-granularity relabeled under-sampling algorithm for imbalanced data. *Applied Soft Computing*, 124. <https://doi.org/10.1016/j.asoc.2022.109083>
- Dangut, M. D., Skaf, Z., & Jennions, I. K. (2022). Handling imbalanced data for aircraft predictive maintenance using the BACHE algorithm. *Applied Soft Computing*, 123. <https://doi.org/10.1016/J.ASOC.2022.108924>
- David Dangut, M., Skaf, Z., & Jennions, I. (2020). Rescaled-LSTM for Predicting Aircraft Component Replacement Under Imbalanced Dataset Constraint. *2020 Advances in Science and Engineering Technology International Conferences (ASET)*. <https://doi.org/10.1109/ASET48392.2020.9118253>
- Dong, Q. L., & Lu, W. (2022). Imbalance Data Classification Method Based on Improved SMOTE Algorithm and Granular Computing. *Cybersecurity and Cyberforensics Conference, 2022-July*, 3196–3201. <https://doi.org/10.23919/CCC55666.2022.9902406>
- Dong, Y., Zhou, C., Wu, Z., & Cheng, M. (2022). Dual Mix-up Adversarial Domain Adaptation for Machine Remaining Useful Life Prediction. *International Conference on Emerging Technologies*, 573–578. <https://doi.org/10.1109/ICET55676.2022.9824464>
- Elkan, C. (2001). The foundations of cost-sensitive learning. *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, 973–978.
- Feng, T., & Zhao, J. (2010). SVM-PHM: A Novel Method for Remaining Useful Life Prediction. *International Conference of Information Science and Management Engineering*, 1, 369–372. <https://doi.org/10.1109/ISME.2010.214>

- Fix, E., & Hodges, J. L. (1989). Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties. *International Statistical Review / Revue Internationale de Statistique*, 57(3), 238. <https://doi.org/10.2307/1403797>
- Forsthoffer, M. S. (2017). Predictive and Preventive Maintenance. *Forsthoffer's More Best Practices for Rotating Equipment*, 501–546. <https://doi.org/10.1016/B978-0-12-809277-4.00011-5>
- Forsthoffer, W. E. (2011). Preventive and Predictive Maintenance Best Practices. *Forsthoffer's Best Practice Handbook for Rotating Machinery*, 563–576. <https://doi.org/10.1016/B978-0-08-096676-2.10011-6>
- Frey, B. J., & Dueck, D. (2007). Clustering by Passing Messages Between Data Points. *Science*, 315(5814), 972–976. <https://doi.org/10.1126/science.1136800>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5). <https://doi.org/10.1214/aos/1013203451>
- Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., & Herrera, F. (2012). A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(4), 463–484. <https://doi.org/10.1109/TSMCC.2011.2161285>
- Ghasemkhani, B., Aktas, O., & Birant, D. (2023). Balanced K-Star: An Explainable Machine Learning Method for Internet-of-Things-Enabled Predictive Maintenance in Manufacturing. *Machines*, 11(3). <https://doi.org/10.3390/MACHINES11030322>
- Ghori, K. M., Abbasi, R. A., Awais, M., Imran, M., Ullah, A., & Szathmary, L. (2020). Performance Analysis of Different Types of Machine Learning Classifiers for Non-Technical Loss Detection. *IEEE Access*, 8, 16033–16048. <https://doi.org/10.1109/ACCESS.2019.2962510>
- Guan, D., Yuan, W., Lee, Y.-K., & Lee, S. (2009). Nearest neighbor editing aided by unlabeled data. *Information Sciences*, 179(13), 2273–2282. <https://doi.org/10.1016/j.ins.2009.02.011>
- Guerroum, M., Zegrari, M., Masmoudi, M., Berquedich, M., & Elmahjoub, A. A. (2022). Machine Learning Technics for Remaining useful Life Prediction using Diagnosis Data: a Case Study of a Jaw Crusher. *International Journal of Emerging Technology and Advanced Engineering*, 12(10), 122–135. [https://doi.org/10.46338/IJETAE1022\\_14](https://doi.org/10.46338/IJETAE1022_14)
- Haibo He, & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>

- Haibo He, Yang Bai, Garcia, E. A., & Shutao Li. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>
- Han, H., Wang, W.-Y., & Mao, B.-H. (2005). *Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning* (pp. 878–887). [https://doi.org/10.1007/11538059\\_91](https://doi.org/10.1007/11538059_91)
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.
- Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and Their Applications*, 13(4), 18–28. <https://doi.org/10.1109/5254.708428>
- Imam, T., Ting, K. M., & Kamruzzaman, J. (2006). *z-SVM: An SVM for Improved Classification of Imbalanced Data* (pp. 264–273). [https://doi.org/10.1007/11941439\\_30](https://doi.org/10.1007/11941439_30)
- Inaba, T. (2021). Performance Evaluation of IoT-enabled Predictive Maintenance. *2021 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2021*, 1551–1555. <https://doi.org/10.1109/IEEM50564.2021.9672828>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning*. Springer US. <https://doi.org/10.1007/978-1-0716-1418-1>
- Jia, Z., Xiao, Z., & Shi, Y. (2021). Remaining Useful Life Prediction of Equipment Based on XGBoost. *International Conference on Computer Science and Application Engineering*. <https://doi.org/10.1145/3487075.3487134>
- Kang, Z., Catal, C., & Tekinerdogan, B. (2021). Remaining Useful Life (RUL) Prediction of Equipment in Production Lines Using Artificial Neural Networks. *Italian National Conference on Sensors*, 21(3), 1–20. <https://doi.org/10.3390/S21030932>
- Kaur, H., Pannu, H. S., & Malhi, A. K. (2020). A Systematic Review on Imbalanced Data Challenges in Machine Learning. *ACM Computing Surveys*, 52(4), 1–36. <https://doi.org/10.1145/3343440>
- Khan, M. M., Tse, P. W., & Trappey, A. J. C. (2021). Development of a Novel Methodology for Remaining Useful Life Prediction of Industrial Slurry Pumps in the Absence of Run to Failure Data. *Italian National Conference on Sensors*, 21(24). <https://doi.org/10.3390/S21248420>
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction.

*Computational and Structural Biotechnology Journal*, 13, 8–17.  
<https://doi.org/10.1016/j.csbj.2014.11.005>

Krzysztof J. Cios, Roman W. Swiniarski, Witold Pedrycz, & Lukasz A. Kurgan. (2007). *Data mining: a knowledge discovery approach*. Springer.

Kubat, M., & Matwin, S. (1997). Addressing The Curse Of Imbalanced Training Sets: One-sided Selection. *International Conference on Machine Learning*, 97, 179.

Lanie, B. L. (2020). Affinity Propagation SMOTE approach for Imbalanced dataset used in Predicting Student at Risk of Low Performance. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(4), 5066–5070.  
<https://doi.org/10.30534/IJATCSE/2020/127942020>

Laureano, L. B., Sison, A. M., & Medina, R. P. (2019). Handling Imbalanced Data through Affinity Propagation and SMOTE. *Proceedings of the 2nd International Conference on Computing and Big Data*, 22–26.  
<https://doi.org/10.1145/3366650.3366665>

Lindgren, T., & Biteus, J. (2017). *APS Failure at Scania Trucks*.  
<https://doi.org/10.24432/C51S51>

Liu, T., Zhu, X., Pedrycz, W., & Li, Z. (2020). A design of information granule-based under-sampling method in imbalanced data classification. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 24(22), 17333–17347. <https://doi.org/10.1007/S00500-020-05023-2>

Ludwig, J. A., & Reynolds, J. F. (1988). *Statistical Ecology: A Primer on Methods and Computing*. Wiley-Interscience Pub.

Mahamad, A. K., Saon, S., & Hiyama, T. (2010). Predicting remaining useful life of rotating machinery based artificial neural network. *Computers and Mathematics with Applications*, 60(4), 1078–1087.  
<https://doi.org/10.1016/J.CAMWA.2010.03.065>

Malhotra, R. (2015). A systematic review of machine learning techniques for software fault prediction. *Applied Soft Computing*, 27, 504–518.  
<https://doi.org/10.1016/j.asoc.2014.11.023>

Matzka, S. (2020). *AI4I 2020 Predictive Maintenance Dataset*.  
<https://doi.org/10.24432/C5HS5C>

Max Bramer. (2007). *Principles of Data Mining*. Springer Science & Business Media.

Mobley, R. K. (2002). Benefits of Predictive Maintenance. *An Introduction to Predictive Maintenance*, 60–73. <https://doi.org/10.1016/B978-075067531-4/50004-X>

Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python A GUIDE FOR DATA SCIENTISTS*.

- Purnachand, K., Shabbeer, M., Syamala Rao, P. N. V. M., & Babu, C. M. (2021). Predictive maintenance of machines and industrial equipment. *Proceedings - 2021 IEEE 10th International Conference on Communication Systems and Network Technologies, CSNT 2021*, 318–324. <https://doi.org/10.1109/CSNT51715.2021.9509696>
- Ronzoni, N., De Marco, A., & Ronchieri, E. (2022). Predictive Maintenance Experiences on Imbalanced Data with Bayesian Optimization Approach. *Communication Systems and Applications, 13377 LNCS*, 120–137. [https://doi.org/10.1007/978-3-031-10536-4\\_9](https://doi.org/10.1007/978-3-031-10536-4_9)
- Sabha, S. U., Assad, A., Din, N. M. U., & Bhat, M. R. (2023). Comparative Analysis of Oversampling Techniques on Small and Imbalanced Datasets Using Deep Learning. *2023 3rd International Conference on Artificial Intelligence and Signal Processing (AISP)*, 1–5. <https://doi.org/10.1109/AISP57993.2023.10134981>
- Sharanya, S., Venkataraman, R., & Murali, G. (2021). Predictive Maintenance : AI Behind Equipment Failure Prediction. *Introduction to AI Techniques for Renewable Energy Systems*, 155–170. <https://doi.org/10.1201/9781003104445-10>
- Shin, J.-H., & Jun, H.-B. (2015). On condition based maintenance policy. *Journal of Computational Design and Engineering*, 2(2), 119–127. <https://doi.org/10.1016/j.jcde.2014.12.006>
- Smith, R., & Mobley, R. K. (2008). Predictive Maintenance Program. *Rules of Thumb for Maintenance and Reliability Engineers*, 47–56. <https://doi.org/10.1016/B978-075067862-9.50005-8>
- Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820. <https://doi.org/10.1109/TII.2014.2349359>
- Tao, X., Li, Q., Ren, C., Guo, W., He, Q., Liu, R., & Zou, J. (2019). Affinity and class probability-based fuzzy support vector machine for imbalanced data sets. *Neural Networks*, 122, 289–307. <https://doi.org/10.1016/J.NEUNET.2019.10.016>
- Tin Kam Ho. (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 278–282. <https://doi.org/10.1109/ICDAR.1995.598994>
- Tomek, I. (1976). Two Modifications of CNN. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(11), 769–772. <https://doi.org/10.1109/TSMC.1976.4309452>
- Unplanned downtime incidents have dropped, but costs continue to rise: MaintainX | Facilities Dive.* (n.d.). Retrieved July 4, 2025, from

<https://www.facilitiesdive.com/news/maintainx-preventive-maintenance-industrial-facilities-management-2024-downtime-costs/724230/>

Wei-Yang Lin, Ya-Han Hu, & Chih-Fong Tsai. (2012). Machine Learning in Financial Crisis Prediction: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(4), 421–436. <https://doi.org/10.1109/TSMCC.2011.2170420>

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47–66. <https://doi.org/10.1016/j.cose.2015.09.005>

Wilson, D. L. (1972). Asymptotic Properties of Nearest Neighbor Rules Using Edited Data. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-2(3)*, 408–421. <https://doi.org/10.1109/TSMC.1972.4309137>

Woods, E., & Kyril, E. (1997). *Ovum evaluates: data mining*. Ovum Ltd.

Xia, S., Zheng, S., Wang, G., Gao, X., & Wang, B. (2021). Granular Ball Sampling for Noisy Label Classification or Imbalanced Classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(4), 2144–2155. <https://doi.org/10.1109/TNNLS.2021.3105984>

Xia, S., Zheng, S., Wang, G., Gao, X., & Wang, B. (2023). Granular Ball Sampling for Noisy Label Classification or Imbalanced Classification. *IEEE Transactions on Neural Networks and Learning Systems*, 34(4), 2144–2155. <https://doi.org/10.1109/TNNLS.2021.3105984>

Xia, T., Shao, Y., Xia, S., Xiong, Y., Lian, X., & Wu, L. (2023). GBSMOTE: A Robust Sampling Method Based on Granular-ball Computing and SMOTE for Class Imbalance. *International Conference on Music and Artificial Intelligence*, 19–24. <https://doi.org/10.1145/3594300.3594304>

Xie, Q., Zhang, Q., Luo, N., & Wang, G. (2024). Three-Way Hybrid Sampling Using Granular Balls for Imbalanced Classification. *IJCRS, 14840 LNAI*, 86–102. [https://doi.org/10.1007/978-3-031-65668-2\\_7](https://doi.org/10.1007/978-3-031-65668-2_7)

Yen, S. J., & Lee, Y. S. (2009). Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, 36(3), 5718–5727. <https://doi.org/10.1016/J.ESWA.2008.06.108>