

DAFTAR PUSTAKA

- Abdelhadi, S., Elbahnasy, K. and Abdelsalam, M. (2020). A proposed model to predict auto insurance claims using machine learning techniques. *Journal of Theoretical and Applied Information Technology*, 98(22), pp.3428–3437.
- Alamir, E., Urgessa, T., Hunegnaw, A. and Gopikrishna, T. (2021). Motor Insurance Claim Status Prediction using Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications*, 12(3).
- Allianz Indonesia. (2020). *Allianz Indonesia Maksimalkan Penggunaan Artificial Intelligence Demi Percepat Proses Klaim Nasabah* [online]. Tersedia dari: <https://www.allianz.co.id/tentang-kami/berita-perusahaan/rilis-media/2020/2020-desember/allianz-indonesia-maksimalkan-penggunaan-artificial-intelligence-demi-percepat-proses-klaim-nasabah.html>
- Anderson, J.F. and Brown, R.L. (2005). *Education and Examination Committee of The Society of Actuaries Risk and Insurance*.
- Aven, T. and Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *Journal of Risk Research*, 12(1), pp.1–11. Tersedia dari: <https://doi.org/10.1080/13669870802488883>.
- Aven, T., Renn, O. and Rosa, E.A. (2011). On the ontological status of the concept of risk. *Safety Science*, 49(8–9).
- Bhagat, R.C. and Patil, S.S. (2015). Enhanced SMOTE algorithm for classification of imbalanced big-data using Random Forest. In: *Souvenir of the 2015 IEEE International Advance Computing Conference, IACC 2015*.
- Bradley, A.P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1).
- Charbuty, B. and Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01).

- Chawla, N. V, Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002). *SMOTE: Synthetic Minority Over-sampling Technique*.
- Chen, T. and Guestrin, C. (2016). XGBoost.
- Chen, X., Wang, M. and Zhang, H. (2011). The use of classification trees for bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1).
- Chinchor, N. (1992). MUC-4 evaluation metrics. In: *4th Message Understanding Conference, MUC 1992 - Proceedings*.
- Conducting Electronic Business Transactions: Why Greater Harmonization Across the Industry is Needed*. (2020). Tersedia dari: <https://www.caqh.org/sites/default/files/explorations/index/report/2019-caqh-index.pdf>
- Cutler, A., Cutler, D.R. and Stevens, J.R. (2012). Random Forests BT - Ensemble Machine Learning: Methods and Applications. In: *Ensemble Machine Learning*.
- Davies, E.R. (2018). Machine learning. In: *Computer Vision*. Elsevier, pp.399–451. Tersedia dari: <https://linkinghub.elsevier.com/retrieve/pii/B9780128092842000149>.
- Dayal, H. (2017). *The Fundamentals of Insurance*. Notion Press.
- Dhieb, N., Ghazzai, H., Besbes, H. and Massoud, Y. (2019). Extreme gradient boosting machine learning algorithm for safe auto insurance operations. In: *2019 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2019*.
- Dimitriadis, S.I. and Liparas, D. (2018). How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: From Alzheimer's disease neuroimaging initiative (ADNI) database. *Neural Regeneration Research*, 13(6), pp.962–970.
- Fauzan, M.A. and Murfi, H. (2018). The accuracy of XGBoost for insurance claim prediction. *International Journal of Advances in Soft Computing and its Applications*, 10(2).

- Ferreira, A.J. and Figueiredo, M.A.T. (2012). Boosting algorithms: A review of methods, theory, and applications. In: *Ensemble Machine Learning: Methods and Applications*.
- Frempong, N.K., Nicholas, N. and Boateng, M.A. (2017). Predictive Modeling, Decision tree, Claim, Risk, Policyholder; Predictive Modeling, Decision tree, Claim, Risk, Policyholder. *International Journal of Statistics and Applications*, 7(2), pp.117–120. Tersedia dari: <http://journal.sapub.org/statistics>.
- Freund, Y. and Schapire, R.E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1).
- Freund, Y. and Schapire, R.E. (1996). Experiments with a New Boosting Algorithm. *Proceedings of the 13th International Conference on Machine Learning*, 1996.
- Friedman, J., Hastie, T. and Tibshirani, R. (2000). *Additive Logistic Regression: A Statistical View of Boosting*.
- Garabaghi, F.H., Benzer, S. and Benzer, R. (2021). Performance Evaluation of Machine Learning Models with Ensemble Learning approach in Classification of Water Quality Indices Based on Different Subset of Features. , 2021. Tersedia dari: <https://doi.org/10.21203/rs.3.rs-876980/v1>.
- Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Second Edition.
- Gonurescu, F. (2011). *Data Mining: Concepts, Model, and Techniques*. Springer.
- Gowda, C.N. and Chaithra, N. (2020). Classification and Regression Tree (CART) Algorithm for the Prediction of Ischemic Heart Disease. *International Journal of Future Generation Communication and Networking*, 13(4), pp.196–206. Tersedia dari: <https://www.researchgate.net/publication/344469894>.
- Han, J., Kamber, M. and Pei, J. (2011). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.

- Hastie, T., Tibshirani, R., James, G. and Witten, D. (2021). An introduction to statistical learning (2nd ed.). *Springer texts*, 102.
- Hastie, T., Tibshirani, R. and Friedman, J. (2009). *Springer Series in Statistics The Elements of Statistical Learning - Data Mining, Inference, and Prediction*.
- Hemalatha, M. and Mahalakshmi, S. (2020). Customer churns prediction in telecom using adaptive logitboost learning approach. *International Journal of Scientific and Technology Research*, 9(2).
- Hemo, S.A. and Rayhan, M.I. (2021). Classification tree and random forest model to predict under-five malnutrition in Bangladesh. *Biom Biostat Int J*, 10(3), pp.116–123.
- Henebry, K.L. and Rejda, G.E. (1995). Principles of Risk Management and Insurance. *The Journal of Risk and Insurance*, 62(4), p.797.
- Iminova, N., Abdimuminova, S. and Mamurov, B. (2019). Risk management and insurance. *International Journal of Innovative Technology and Exploring Engineering*, 8(12).
- Indonesia. (2014). Undang-Undang Republik Indonesia Nomor 40 Tahun 2014 tentang Perasuransian.
- Innes, C. (2020). Claims analytics. An actuarial and data scientist perspective. In: *Society of Actuaries*. Society of Actuaries. Tersedia dari: <https://www.soa.org/4a6b09/globalassets/assets/files/e-business/pd/events/2020/predictive-analytics-4-0/pd-2020-09-pas-session-013.pdf>.
- Jordan, M.I. and Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 2015.
- Kamarudin, M.H., Maple, C., Watson, T. and Safa, N. (2017). A LogitBoost-Based Algorithm for Detecting Known and Unknown Web Attacks. *IEEE Access*, PP, p.1.
- Kaur, H. and Kumari, V. (2022). Predictive modelling and analytics for diabetes using a machine learning approach. *Applied Computing and Informatics*, 18(1–2), pp.90–100.

- Kumar, P., Bhatnagar, R., Gaur, K. and Bhatnagar, A. (2021). Classification of Imbalanced Data: Review of Methods and Applications. *IOP Conference Series: Materials Science and Engineering*, 1099(1), p.012077.
- Kusumastuti, A. (2022). *Implementasi Extreme Gradient Boosting (XGBoost) dalam Menangani Imbalanced Class pada Analisis Klasifikasi* [unpublished]. Yogyakarta: Universitas Gadjah Mada.
- Lemos, F. (2020). On the definition of risk. *Journal of Risk Management in Financial Institutions*, 13(3).
- Li, P. (2010). Robust LogitBoost and Adaptive Base Class (ABC) LogitBoost. In: *The Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*.
- Lin, W., Wu, Z., Lin, L., Wen, A. and Li, J. (2017). An ensemble random forest algorithm for insurance big data analysis. *IEEE Access*, 5.
- Loh, W.Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), pp.14–23.
- Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *BBA - Protein Structure*, 405(2).
- Mavrin, A. (2019). *LogitBoost* [online]. Tersedia dari: <https://logitboost.readthedocs.io/>
- McDonald, R.A., Hand, D.J. and Eckley, I.A. (2003). An empirical comparison of three boosting algorithms on real data sets with artificial class noise. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2709.
- McDonald, R.A., Hand, J. and Eckley, I.A. *An Empirical Comparison of Three Boosting Algorithms on Real Data Sets with Artificial Class Noise*.
- Michael, J.M. and E.Rejda, G. (2017). *Principles of Risk Management and Insurance [ebook]*.
- Müller, A.C. and Guido, S. (2017). *Introduction to Machine Learning with Python: a guide for data scientist*.

- Mumford, C.L. (2009). *Computational intelligence : collaboration, fusion and emergence*. Springer.
- Outreville, J.F. (1998). Insurance Concepts. In: *Theory and Practice of Insurance*. Springer US, pp.131–146.
- Paltrinieri, N., Comfort, L. and Reniers, G. (2019). Learning about risk: Machine learning for risk assessment. *Safety Science*, 118.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel V. and Thirion, B., Grisel, O., Blondel, M., Prettenhofer P. and Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, pp.2825–2830.
- Pesantez-Narvaez, J., Guillen, M. and Alcañiz, M. (2019). Predicting motor insurance claims using telematics data—XGboost versus logistic regression. *Risks*, 7(2).
- Prayoga, R. (2020). *Analisis Perbandingan Random Forest, AdaBoost, dan Gradient Tree Boosting pada Imbalanced Dataset* [unpublished]. Tesis, Yogyakarta: Universitas Gadjah Mada.
- Rausand, M. and Haugen, S. (2020). *Risk assessment: Theory, methods, and applications*.
- Rejda, G.E. and McNamara, M.J. (2014). *Principles of Risk Management and Insurance*. 12th ed. Pearson.
- Rowe, W.D. (1994). Understanding Uncertainty. *Risk Analysis*, 14(5).
- Rusdah, D.A. and Murfi, H. (2020). XGBoost in handling missing values for life insurance risk prediction. *SN Applied Sciences*, 2(8).
- Salem, S. Ben, Naouali, S. and Sallami, M. (2017). Clustering Categorical Data Using the K-Means Algorithm and the Attribute's Relative Frequency. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 11, pp.708–713. Tersedia dari: <https://api.semanticscholar.org/CorpusID:54042483>.

- Samuel, A.L. (1959). *Some Studies in Machine Learning Using the Game of Checkers*.
- Schapire, R.E. (1999). A brief introduction to boosting. In: *IJCAI International Joint Conference on Artificial Intelligence*.
- Schapire, R.E. (1990). The strength of weak learnability. *Machine Learning*, 5(2).
- Schonlau, M. and Zou, R.Y. (2020). The random forest algorithm for statistical learning. *Stata Journal*, 20(1), pp.3–29.
- Shalev-Shwartz, S. and Ben-David, S. (2013). *Understanding machine learning: From theory to algorithms*.
- Song, K., Yan, F., Ding, T., Gao, L. and Lu, S. (2020). A steel property optimization model based on the XGBoost algorithm and improved PSO. *Computational Materials Science*, 174.
- Suthaharan, S. (2015). *Machine Learning Models and Algorithms for Big Data Classification*. Boston, MA: Springer US. Tersedia dari: <https://link.springer.com/10.1007/978-1-4899-7641-3>.
- Tampubulon, B. (2022). Naik Signifikan, Penetrasi Asuransi Jiwa Tembus 8%. In: *Asosiasi Asuransi Jiwa Indonesia*. Tersedia dari: <https://aaji.or.id/file/uploads/content/file/1.%20Siaran%20Pers%20Semester%201%202022%20-%20rev.pdf>
- Tan, A.C. and Gilbert, D. (2003). An empirical comparison of supervised machine learning techniques in bioinformatics. *Proceedings of the First Asia-Pacific bioinformatics conference on Bioinformatics 2003 - Volume 19*, 19.
- Thoyts, R. (2010). *Insurance Theory and Practice*.
- Wardana, A.P. (2020). *Klasifikasi dan Pengelompokan Polis Asuransi Kendaraan Bermotor: Upaya Persiapan Implementasi IFRS 17 pada Perusahaan Asuransi Umum* [unpublished]. Yogyakarta: Universitas Gadjah Mada.
- Weerasinghe, K.P.M.L.P. and Wijegunasekara, M.C. (2016). A Comparative Study of Data Mining Algorithms in the Prediction of Auto Insurance Claims. *European International Journal of Science and Technology*, 5(1).

- Williams, G. (2011). *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*.
- Wirawan, D. (2022). *Penentuan Premi Murni Asuransi Kendaraan Bermotor berdasarkan Jarak Tempuh (Pay-As-You-Drive Insurance) dengan Tree-Based Machine Learning* [unpublished]. Skripsi, Yogyakarta: Universitas Gadjah Mada.
- World Health Organization. (2004). *ICD-10 Version:2019*. Tersedia dari: <https://icd.who.int/browse10/2019/en>
- XGBoost Developers. (2022). *XGBoost Documentation* [online]. Tersedia dari: <https://xgboost.readthedocs.io/en/stable/index.html>.
- Zhang, W., Wu, C., Zhong, H., Li, Y. and Wang, L. (2021). Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geoscience Frontiers*, 12(1).
- Zhu, R., Guo, Y. and Xue, J.H. (2020). Adjusting the imbalance ratio by the dimensionality of imbalanced data. *Pattern Recognition Letters*, 133.