DAFTAR PUSTAKA

Abhilasha Narote, Abhijeet Pawar, Mansi Gaikwad, Tanuja Dalvi, & Purva Gondhalekar. (2022). Prediction of Diabetes Mellitus using Machine Learning. *International Journal of Advanced Research in Science, Communication and Technology*, *10*(04), 17–20. https://doi.org/10.48175/ijarsct-2988

Adem, K. (2020). Diagnosis of breast cancer with Stacked autoencoder and Subspace kNN. *Physica A: Statistical Mechanics and Its Applications*, *551*, 124591. https://doi.org/10.1016/j.physa.2020.124591

Aggarwal, C. C., Kong, X., Gu, Q., Han, J., & Yu, P. S. (2014). Active learning: A survey. *Data Classification: Algorithms and Applications*, 571–605. https://doi.org/10.1201/b17320

Ali, H., Salleh, M. N. M., Saedudin, R., Hussain, K., & Mushtaq, M. F. (2019). Imbalance class problems in data mining: A review. *Indonesian Journal of Electrical Engineering and Computer Science*, *14*(3), 1552–1563. https://doi.org/10.11591/ijeecs.v14.i3.pp1552-1563

Alirezanejad, M., Enayatifar, R., Motameni, H., & Nematzadeh, H. (2020). Heuristic filter feature selection methods for medical datasets. *Genomics*, *112*(2), 1173–1181. https://doi.org/10.1016/j.ygeno.2019.07.002

Alkhasawneh, M. S. (2022). Software Defect Prediction through Neural Network and Feature Selections. *Applied Computational Intelligence and Soft Computing*, *2022*, 1–16. https://doi.org/10.1155/2022/2581832

Angel Viji, K. S., & Hevin Rajesh, D. (2019). An Efficient Technique to Segment the Tumor and Abnormality Detection in the Brain MRI Images Using KNN Classifier. *Materials Today: Proceedings*, *24*, 1944–1954. https://doi.org/10.1016/j.matpr.2020.03.622

Arslan, H., & Arslan, H. (2021). A new COVID-19 detection method from human genome sequences using CpG island features and KNN classifier. *Engineering Science and Technology, an International Journal*, *24*(4), 839–847. https://doi.org/10.1016/j.jestch.2020.12.026

Azuaje, F. (2006). Witten IH, Frank E: Data Mining: Practical Machine Learning Tools and Techniques 2nd edition. *BioMedical Engineering OnLine*, *5*(1). https://doi.org/10.1186/1475-925x-5-51

Bagui, S. C. (2005). Combining Pattern Classifiers: Methods and Algorithms. In *Technometrics* (Vol. 47, Issue 4). https://doi.org/10.1198/tech.2005.s320

Bakri, R., Astuti, N. P., & Ahmar, A. S. (2022). Machine Learning Algorithms with Parameter Tuning to Predict Students' Graduation-on-time: A Case Study in Higher Education. *Journal of Applied Science, Engineering, Technology, and Education*, *4*(2), 259–265. https://doi.org/10.35877/454ri.asci1581

Barua, S., Islam, M. M., Yao, X., & Murase, K. (2014). MWMOTE - Majority weighted minority oversampling technique for imbalanced data set learning. *IEEE Transactions on Knowledge and Data Engineering*, *26*(2), 405–425. https://doi.org/10.1109/TKDE.2012.232

Bashir, S., Khattak, I. U., Khan, A., Khan, F. H., Gani, A., & Shiraz, M. (2022). A Novel Feature Selection Method for Classification of Medical Data Using Filters, Wrappers, and Embedded Approaches. *Complexity*, *2022*. https://doi.org/10.1155/2022/8190814

Basu, A., Roy, R., & Savitha, N. (2018). Performance Analysis of Regression and Classification Models in the Prediction of Breast Cancer. *Indian Journal of Science and Technology*, *11*(3), 1–6. https://doi.org/10.17485/ijst/2018/v11i3/119179

Beckmann, M., Ebecken, N. F. F., & Pires de Lima, B. S. L. (2015). A KNN Undersampling Approach for Data Balancing. *Journal of Intelligent Learning Systems and Applications*, *07*(04), 104–116. https://doi.org/10.4236/jilsa.2015.74010

Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation Measures for Models Assessment over Imbalanced Data Sets. *Journal of Information Engineering and Applications*, *3*(10), 27–38. http://www.iiste.org/Journals/index.php/JIEA/article/view/7633

Bennett, R., Mulla, Z. D., Parikh, P., Hauspurg, A., & Razzaghi, T. (2022). An imbalance-aware deep neural network for early prediction of preeclampsia. In *PLoS ONE* (Vol. 17, Issue 4 April). https://doi.org/10.1371/journal.pone.0266042

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011*, 1–9.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A. L., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *13*(2), 1–43. https://doi.org/10.1002/widm.1484

Borowska, K., & Stepaniuk, J. (2016). Imbalanced data classification: A novel re-sampling approach combining versatile improved SMOTE and rough sets. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9842 LNCS*(1), 31–42. https://doi.org/10.1007/978-3-319-45378-1_4

Bunkhumpornpat, C., Sinapiromsaran, K., & Lursinsap, C. (2009). Safe-level-SMOTE: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. *Lecture Notes in Computer Science (Including*

Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), *5476 LNAI*, 475–482. https://doi.org/10.1007/978-3-642-01307-2_43

Carreno, J. F., & Qiu, P. (2020). Feature selection algorithms for predicting preeclampsia: A comparative approach. *Proceedings - 2020 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2020*, 2626–2631. https://doi.org/10.1109/BIBM49941.2020.9313108

Chawla, N. V., Japkowicz, N., & Kotcz, A. (2004). Editorial. *ACM SIGKDD Explorations Newsletter*, *6*(1), 1–6. https://doi.org/10.1145/1007730.1007733

Chen, H., Li, T., Fan, X., & Luo, C. (2019). Feature selection for imbalanced data based on neighborhood rough sets. *Information Sciences*, *483*, 1–20. https://doi.org/10.1016/j.ins.2019.01.041

Chen, Z., Duan, J., Kang, L., & Qiu, G. (2021). A hybrid data-level ensemble to enable learning from highly imbalanced dataset. *Information Sciences*, *554*, 157–176. https://doi.org/10.1016/j.ins.2020.12.023

Chopra, D., & Khurana, R. (2023). Introduction to Machine Learning with Python. In *Introduction to Machine Learning with Python*. https://doi.org/10.2174/97898151244221230101

Datta, P., Das, P., & Kumar, A. (2022). Hyper parameter tuning based gradient boosting algorithm for detection of diabetic retinopathy: an analytical review. *Bulletin of Electrical Engineering and Informatics*, *11*(2), 814–824. https://doi.org/10.11591/eei.v11i2.3559

Dimitriadis, E., Rolnik, D. L., Zhou, W., Estrada-Gutierrez, G., Koga, K., Francisco, R. P. V., Whitehead, C., Hyett, J., da Silva Costa, F., Nicolaides, K., & Menkhorst, E. (2023). Pre-eclampsia. *Nature Reviews Disease Primers*, *9*(1), 1–22. https://doi.org/10.1038/s41572-023-00417-6

Ding, H., Chen, L., Dong, L., Fu, Z., & Cui, X. (2022). Imbalanced data classification: A KNN and generative adversarial networks-based hybrid approach for intrusion detection. *Future Generation Computer Systems*, *131*, 240–254. https://doi.org/10.1016/j.future.2022.01.026

Douzas, G., Bacao, F., & Last, F. (2018). Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Information Sciences*, *465*, 1–20. https://doi.org/10.1016/j.ins.2018.06.056

Drosou, K., & Koukouvinos, C. (2017). Proximal support vector machine techniques on medical prediction outcome. *Journal of Applied Statistics*, *44*(3), 533–553. https://doi.org/10.1080/02664763.2016.1177499

Farzaneh, F., Tavakolikia, Z., & Soleimanzadeh Mousavi, S. H. (2019). Assessment of occurrence of preeclampsia and some clinical and demographic risk factors in Zahedan city in 2017. *Clinical and Experimental Hypertension*, *41*(6), 583–588. https://doi.org/10.1080/10641963.2018.1523919

Fox, R., Kitt, J., Leeson, P., Aye, C. Y. L., & Lewandowski, A. J. (2019). *Preeclampsia : Risk Factors , Diagnosis , Management , and the Cardiovascular Impact on the O ff spring*. 1–22.

Fuadah, Y. N., Pramudito, M. A., & Lim, K. M. (2023). An Optimal Approach for Heart Sound Classification Using Grid Search in Hyperparameter Optimization of Machine Learning. *Bioengineering*, *10*(1). https://doi.org/10.3390/bioengineering10010045

Gan, D., Shen, J., An, B., Xu, M., & Liu, N. (2020). Integrating TANBN with cost sensitive classification algorithm for imbalanced data in medical diagnosis. *Computers and Industrial Engineering*, *140*(December 2019), 106266. https://doi.org/10.1016/j.cie.2019.106266

García-García, J. C., & García-Ródenas, R. (2021). A methodology for automatic parameter-tuning and center selection in density-peak clustering methods. *Soft Computing*, *25*(2), 1543–1561. https://doi.org/10.1007/s00500-020-05244-5

García, V., Sánchez, J. S., & Mollineda, R. A. (2012). On the effectiveness of preprocessing methods when dealing with different levels of class imbalance. *Knowledge-Based Systems*, *25*(1), 13–21. https://doi.org/10.1016/j.knosys.2011.06.013

Gbenga, D. E., Christopher, N., & Yetunde, D. C. (2017). Performance Comparison of Machine Learning Techniques for Breast Cancer Detection. *Nova Journal of Engineering and Applied Sciences*, *6*(1), 1–8. https://doi.org/10.20286/nova-jeas-060105

Géron, A. (2019). Hands-on Machine Learning whith Scikit-Learing, Keras and Tensorfow. In *O'Reilly Media, Inc*.

Gorunescu, F. (2011). *Data Mining - Intelligent Systems Reference Library*. 364. https://link.springer.com/book/10.1007/978-3-642-19721-5

Guido, R., Groccia, M. C., & Conforti, D. (2022). A hyper-parameter tuning approach for cost-sensitive support vector machine classifiers. *Soft Computing*, *27*(18), 12863–12881. https://doi.org/10.1007/s00500-022-06768-8

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, *73*, 220–239. https://doi.org/10.1016/j.eswa.2016.12.035

Haixiang, G., Yijing, L., Yanan, L., Xiao, L., & Jinling, L. (2016). BPSO-Adaboost-KNN ensemble learning algorithm for multi-class imbalanced data classification. *Engineering Applications of Artificial Intelligence*, *49*, 176–193. https://doi.org/10.1016/j.engappai.2015.09.011

Han, H., Wang, W., & Mao, B. (2005). *Borderline-SMOTE : A New Over-Sampling Method in*. 878–887.

Hartono, & Ongko, E. (2022). Avoiding Overfitting dan Overlapping in Handling

Class Imbalanced Using Hybrid Approach with Smoothed Bootstrap Resampling and Feature Selection. *International Journal on Informatics Visualization*, *6*(2), 343–348. https://doi.org/10.30630/joiv.6.2.985

He, Haibo, & 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

He, Hongliang, Zhang, W., & Zhang, S. (2018). A novel ensemble method for credit scoring: Adaption of different imbalance ratios. *Expert Systems with Applications*, *98*, 105–117. https://doi.org/10.1016/j.eswa.2018.01.012

Hossain, R., & Timmer, D. D. (2021). Machine learning model optimization with hyper parameter tuning approach. *Global Journal of Computer Science and Technology*, *21*(2), 7–13.

Hussein, A. S., Li, T., Yohannese, C. W., & Bashir, K. (2019). A-SMOTE: A new preprocessing approach for highly imbalanced datasets by improving SMOTE. *International Journal of Computational Intelligence Systems*, *12*(2), 1412–1422. https://doi.org/10.2991/ijcis.d.191114.002

Islam, A., Belhaouari, S. B., Rehman, A. U., & Bensmail, H. (2022). KNNOR: An oversampling technique for imbalanced datasets[Formula presented]. *Applied Soft Computing*, *115*, 108288. https://doi.org/10.1016/j.asoc.2021.108288

Jain, D., & Singh, V. (2018a). Feature selection and classification systems for chronic disease prediction : A review. *Egyptian Informatics Journal*, *19*(3), 179–189. https://doi.org/10.1016/j.eij.2018.03.002

Jain, D., & Singh, V. (2018b). Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal*, *19*(3), 179–189. https://doi.org/10.1016/j.eij.2018.03.002

Jhee, J. H., Lee, S., Park, Y., Lee, S. E., Kim, Y. A., Kang, W., Kwon, J., & Park, J. T. (2019). *Prediction model development of late-onset preeclampsia using machine learning-based methods*. 1–12.

Jian, C., Gao, J., & Ao, Y. (2016). A new sampling method for classifying imbalanced data based on support vector machine ensemble. *Neurocomputing*, *193*, 115–122. https://doi.org/10.1016/j.neucom.2016.02.006

Jin, H. (2022). *Hyperparameter Importance for Machine Learning Algorithms*. 1–8. http://arxiv.org/abs/2201.05132

Kaur, H., Pannu, H. S., & Malhi, A. K. (2019). A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Computing Surveys*, *52*(4). https://doi.org/10.1145/3343440

Khaldy, M. Al, & Kambhampati, C. (2018). *Resampling imbalanced class and the effectiveness of feature selection methods for heart failure dataset*. *4*(1), 37–45. https://doi.org/10.15406/iratj.2018.04.00090

Khan, S., Khan, A., Maqsood, M., Aadil, F., & Ghazanfar, M. A. (2019). Optimized Gabor Feature Extraction for Mass Classification Using Cuckoo Search for Big Data E-Healthcare. *Journal of Grid Computing*, *17*(2), 239–254. https://doi.org/10.1007/s10723-018-9459-x

Kong, J., Kowalczyk, W., Nguyen, D. A., Back, T., & Menzel, S. (2019). Hyperparameter Optimisation for Improving Classification under Class Imbalance. *2019 IEEE Symposium Series on Computational Intelligence, SSCI 2019*, 3072–3078. https://doi.org/10.1109/SSCI44817.2019.9002679

Kraiem, M. S., Sánchez-Hernández, F., & Moreno-García, M. N. (2021). Selecting the suitable resampling strategy for imbalanced data classification regarding dataset properties. An approach based on association models. *Applied Sciences (Switzerland)*, *11*(18). https://doi.org/10.3390/app11188546

Landset, S., Khoshgoftaar, T. M., Richter, A. N., & Hasanin, T. (2015). A survey of open source tools for machine learning with big data in the Hadoop ecosystem. *Journal of Big Data*, 1–36. https://doi.org/10.1186/s40537-015-0032-1

Li, J., Zhu, Q., & Wu, Q. (2020). A parameter-free hybrid instance selection algorithm based on local sets with natural neighbors. *Applied Intelligence*, *50*(5), 1527–1541. https://doi.org/10.1007/s10489-019-01598-y

Li, J., Zhu, Q., Wu, Q., & Fan, Z. (2021). A novel oversampling technique for class-imbalanced learning based on SMOTE and natural neighbors. *Information Sciences*, *565*, 438–455. https://doi.org/10.1016/j.ins.2021.03.041

Li, Y. xin, Shen, X. ping, Yang, C., Cao, Z. zeng, Du, R., Yu, M. da, Wang, J. ping, & Wang, M. (2021). Novel electronic health records applied for prediction of pre-eclampsia: Machine-learning algorithms. *Pregnancy Hypertension*, *26*(August), 102–109. https://doi.org/10.1016/j.preghy.2021.10.006

Lin, Y. C., Mallia, D., Clark-Sevilla, A. O., Catto, A., Leshchenko, A., Haas, D. M., Wapner, R., Pe'er, I., Raja, A., & Salleb-Aouissi, A. (2022). Preeclampsia Predictor with Machine Learning: A Comprehensive and Bias-Free Machine Learning Pipeline. *MedRxiv*, 2022.06.08.22276107.

Liu, J., & Zio, E. (2019). Integration of feature vector selection and support vector machine for classification of imbalanced data. *Applied Soft Computing Journal*, *75*, 702–711. https://doi.org/10.1016/j.asoc.2018.11.045

Liu, M., Yang, X., Chen, G., Ding, Y., Shi, M., Sun, L., Huang, Z., Liu, J., Liu, T., Yan, R., & Li, R. (2022). Development of a prediction model on preeclampsia using machine learning-based method: a retrospective cohort study in China. *Frontiers in Physiology*, *13*(August), 1–9. https://doi.org/10.3389/fphys.2022.896969

Liu, N., Li, X., Qi, E., Xu, M., Li, L., & Gao, B. (2020). A novel ensemble learning paradigm for medical diagnosis with imbalanced data. *IEEE Access*, *8*, 171263–171280. https://doi.org/10.1109/ACCESS.2020.3014362

Macarena Espinilla, 1 Javier Medina, 1 Ángel-Luis García-Fernández, 1 Sixto Campaña, 2 and Jorge Londoño3. (2017). Fuzzy Intelligent System for Patients with Preeclampsia in Wearable Devices. *Mobile Information Systems*, *2017*. https://doi.org/10.1155/2017/7838464

Mahfouz, M. A., Shoukry, A., & Ismail, M. A. (2021). EKNN: Ensemble classifier incorporating connectivity and density into kNN with application to cancer diagnosis. *Artificial Intelligence in Medicine*, *111*(December 2019), 101985. https://doi.org/10.1016/j.artmed.2020.101985

Manoochehri, Z., Manoochehri, S., Soltani, F., Tapak, L., & Sadeghifar, M. (2021). Predicting preeclampsia and related risk factors using data mining approaches: A cross-sectional study. *International Journal of Reproductive BioMedicine*, *19*(11), 959–968. https://doi.org/10.18502/ijrm.v19i11.9911

Mantovani, R. G., Horvath, T., Cerri, R., Vanschoren, J., & De Carvalho, A. C. P. L. F. (2017). Hyper-Parameter Tuning of a Decision Tree Induction Algorithm. *Proceedings - 2016 5th Brazilian Conference on Intelligent Systems, BRACIS 2016*, 37–42. https://doi.org/10.1109/BRACIS.2016.018

Marić, I., Tsur, A., Aghaeepour, N., Montanari, A., Stevenson, D. K., Shaw, G. M., & Winn, V. D. (2020). Early prediction of preeclampsia via machine learning. *American Journal of Obstetrics and Gynecology MFM*, *2*(2), 100100. https://doi.org/10.1016/j.ajogmf.2020.100100

Marin, I., Pavaloiu, B. I., Marian, C. V., Racovita, V., & Goga, N. (2019). Early detection of preeclampsia based on a machine learning approach. *2019 7th E-Health and Bioengineering Conference, EHB 2019*, 23–26. https://doi.org/10.1109/EHB47216.2019.8970025

Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., & De Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Research Notes*, *4*(1), 299. https://doi.org/10.1186/1756-0500-4-299

Martel, A. (2023). *Distance rank score: unsupervised filter method for feature selection on imbalanced dataset*. 1–14.

Mayrink, J., Costa, M. L., & Cecatti, J. G. (2018). Preeclampsia in 2018: Revisiting Concepts, Physiopathology, and Prediction. *Scientific World Journal*, *2018*. https://doi.org/10.1155/2018/6268276

Mera-Gaona, M., López, D. M., Vargas-Canas, R., & Neumann, U. (2021). Framework for the ensemble of feature selection methods. *Applied Sciences (Switzerland)*, *11*(17), 1–16. https://doi.org/10.3390/app11178122

Mienye, I. D., & Sun, Y. (2021). Performance analysis of cost-sensitive learning methods with application to imbalanced medical data. *Informatics in Medicine Unlocked*, *25*, 100690. https://doi.org/10.1016/j.imu.2021.100690

Mikat, B., Gellhaus, A., Wagner, N., Birdir, C., & Kimmig, R. (2012). *Early Detection of Maternal Risk for Preeclampsia*. *2012*. https://doi.org/10.5402/2012/172808

Modak, R., Pal, A., Pal, A., & Ghosh, M. K. (2020). Prediction of preeclampsia by a combination of maternal spot urinary protein-creatinine ratio and uterine artery doppler. *International Journal of Reproduction, Contraception, Obstetrics and Gynecology*, *9*(2), 635. https://doi.org/10.18203/2320-1770.ijrcog20200350

Moeloek, A., Lampung, P., Januari, P., Aulia, D., Graharti, R., Kedokteran, F., Lampung, U., Obstetrik, B., Kedokteran, F., Lampung, U., Klinik, B. P., Kedokteran, F., & Lampung, U. (2019). *Hubungan Diabetes Melitus dengan Kejadian Preeklampsia di RSUD DR . H . Relationship between Diabetes Mellitus with The Incidence of Preeclampsia in RSUD DR . H . Abdul Moeloek Lampung on The Period*. *8*, 180–186.

Moghadas, E., Rezazadeh, J., & Farahbakhsh, R. (2020). An IoT patient monitoring based on fog computing and data mining: Cardiac arrhythmia usecase. *Internet of Things (Netherlands)*, *11*. https://doi.org/10.1016/j.iot.2020.100251

Morais, R. F. A. B. De, & Vasconcelos, G. C. (2017). *Under-Sampling the Minority Class to Improve the Performance of Over- Sampling Algorithms in Imbalanced Data Sets Under-Sampling the Minority Class to Improve the Performance of Over-Sampling Algorithms in Imbalanced Data Sets*. *September*.

Moreira, M. W. L., Rodrigues, J. J. P. C., Oliveira, A. M. B., Ramos, R. F., & Saleem, K. (2016). A preeclampsia diagnosis approach using Bayesian networks. *2016 IEEE International Conference on Communications, ICC 2016*. https://doi.org/10.1109/ICC.2016.7510893

Moreno-Ibarra, M. A., Villuendas-Rey, Y., Lytras, M. D., Yáñez-Márquez, C., & Salgado-Ramírez, J. C. (2021). Classification of diseases using machine learning algorithms: A comparative study. *Mathematics*, *9*(15), 1–21. https://doi.org/10.3390/math9151817

Mueni, F., & Mwangi, M. (2019). *Informatics in Medicine Unlocked A 24-hour ambulatory blood pressure monitoring system for preeclampsia management in antenatal care*. *16*(May). https://doi.org/10.1016/j.imu.2019.100199

Musyoka, F. M., Thiga, M. M., & Muketha, G. M. (2019). A 24-hour ambulatory blood pressure monitoring system for preeclampsia management in antenatal care. *Informatics in Medicine Unlocked*, *16*(June). https://doi.org/10.1016/j.imu.2019.100199

Nagarajan, S. M., Muthukumaran, V., Murugesan, R., Joseph, R. B., & Munirathanam, M. (2021). Feature selection model for healthcare analysis and classification using classifier ensemble technique. *International Journal of Systems Assurance Engineering and Management*.

https://doi.org/10.1007/s13198-021-01126-7

Nair, P., & Kashyap, I. (2019). Optimization of kNN classifier using hybrid preprocessing model for handling imbalanced data. *International Journal of Engineering Research and Technology*, *12*(5), 697–704.

Navaratnam, K., Alfirevic, Z., Baker, P. N., Gluud, C., Grüttner, B., Kublickiene, K., Zeeman, G., & Kenny, L. C. (2013). A multi-centre phase IIa clinical study of predictive testing for preeclampsia: Improved pregnancy outcomes via early detection (IMPROvED). *BMC Pregnancy and Childbirth*, *13*, 1–7. https://doi.org/10.1186/1471-2393-13-226

Nwe, M. M., & Lynn, K. T. (2020). KNN-Based Overlapping Samples Filter Approach for Classification of Imbalanced Data. In *Studies in Computational Intelligence* (Vol. 845). Springer International Publishing. https://doi.org/10.1007/978-3-030-24344-9_4

Obstetri, P. (2016). *PRE-EKLAMSIA*.

Pan, Z., Wang, Y., & Ku, W. (2017). A new k-harmonic nearest neighbor classifier based on the multi-local means. *Expert Systems with Applications*, *67*, 115–125. https://doi.org/10.1016/j.eswa.2016.09.031

Panda, B. (2020). *A survey on application of Population Based Algorithm on Hyperparameter Selection*. *April*. https://doi.org/10.13140/RG.2.2.11820.21128

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Piri, S., Delen, D., & Liu, T. (2018). A synthetic informative minority over-sampling (SIMO) algorithm leveraging support vector machine to enhance learning from imbalanced datasets. *Decision Support Systems*, *106*(April 2023), 15–29. https://doi.org/10.1016/j.dss.2017.11.006

Rajesh, K. N. V. P. S., & Dhuli, R. (2018). Classification of imbalanced ECG beats using re-sampling techniques and AdaBoost ensemble classifier. *Biomedical Signal Processing and Control*, *41*, 242–254. https://doi.org/10.1016/j.bspc.2017.12.004

Ramos-pérez, I., Arnaiz-gonzález, Á., Rodríguez, J. J., & García-osorio, C. (2022). *When is resampling beneficial for feature selection with imbalanced wide data ? 188*.

Rana, S., Lemoine, E., Granger, J., & Karumanchi, S. A. (2019). Preeclampsia: Pathophysiology, Challenges, and Perspectives. *Circulation Research*, *124*(7), 1094–1112. https://doi.org/10.1161/CIRCRESAHA.118.313276

Rekha, G., Tyagi, A. K., & Reddy, V. K. (2019). A wide scale classification of class imbalance problem and its solutions: A systematic literature review. *Journal of Computer Science*, *15*(7), 886–929. https://doi.org/10.3844/jcssp.2019.886.929

Rezende, K. B. de C., Cunha, A. J. L. A. da, Pritsivelis, C., Faleiro, E. C., Amim Junior, J., & Bornia, R. G. (2019). How do maternal factors impact preeclampsia prediction in Brazilian population? *Journal of Maternal-Fetal and Neonatal Medicine*, *32*(7), 1051–1056. https://doi.org/10.1080/14767058.2017.1399115

Rokotyanskaya, E. A., Panova, I. A., Malyshkina, A. I., Fetisova, I. N., Fetisov, N. S., Kharlamova, N. V., & Kuligina, M. V. (2020). Technologies for prediction of preeclampsia. *Sovremennye Tehnologii v Medicine*, *12*(5), 78–86. https://doi.org/10.17691/stm2020.12.5.09

Sáez, J. A., Luengo, J., Stefanowski, J., & Herrera, F. (2015). SMOTE-IPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. *Information Sciences*, *291*(C), 184–203. https://doi.org/10.1016/j.ins.2014.08.051

Saranya, G., & Pravin, A. (2024). Grid Search based Optimum Feature Selection by Tuning hyperparameters for Heart Disease Diagnosis in Machine learning. *The Open Biomedical Engineering Journal*, *17*(1), 1–13. https://doi.org/10.2174/18741207-v17-e230510-2022-ht28-4371-8

Serra, B., Mendoza, M., Scazzocchio, E., Meler, E., Nolla, M., Sabrià, E., Rodríguez, I., & Carreras, E. (2020). A new model for screening for early-onset preeclampsia. *American Journal of Obstetrics and Gynecology*, *222*(6), 608.e1-608.e18. https://doi.org/10.1016/j.ajog.2020.01.020

Shaban, W. M., Rabie, A. H., Saleh, A. I., & Abo-Elsoud, M. A. (2020). A new COVID-19 Patients Detection Strategy (CPDS) based on hybrid feature selection and enhanced KNN classifier. *Knowledge-Based Systems*, *205*, 106270. https://doi.org/10.1016/j.knosys.2020.106270

Shardlow, M. (2016). An Analysis of Feature Selection Techniques. *The University of Manchester*, *14*(1), 1–7.

Shi, Z. (2020). Improving k-Nearest Neighbors Algorithm for Imbalanced Data Classification. *IOP Conference Series: Materials Science and Engineering*, *719*(1). https://doi.org/10.1088/1757-899X/719/1/012072

Simbolon, O., Widyawati, M. N., Kurnianingsih, K., Kubota, N., & Ng, N. (2020). Predicting the Risk of Preeclampsia using Soft Voting-based Ensemble and Its Recommendation. *2020 International Symposium on Community-Centric Systems, CcS 2020*. https://doi.org/10.1109/CcS49175.2020.9231400

Singh, N. D., & Dhall, A. (2018). *Clustering and Learning from Imbalanced Data*. 1–9. http://arxiv.org/abs/1811.00972

Soongsatitanon, A., & Phupong, V. (2022). Prediction of preeclampsia using first trimester placental protein 13 and uterine artery Doppler. *Journal of Maternal-Fetal and Neonatal Medicine*, *35*(22), 4412–4417. https://doi.org/10.1080/14767058.2020.1849127

Sufriyana, H., Wu, Y. W., & Su, E. C. Y. (2020). Artificial intelligence-assisted prediction of preeclampsia: Development and external validation of a nationwide health insurance dataset of the BPJS Kesehatan in Indonesia. *EBioMedicine*, *54*. https://doi.org/10.1016/j.ebiom.2020.102710

Sun, Y., Wong, A. K. C., & Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, *23*(4), 687–719. https://doi.org/10.1142/S0218001409007326

Suprihatin, E., & Wuryaningsih, S. H. (2019). *Prediksi Preeklampsi Secara Dini Melalui Pengukuran Body Mass Index , Mean Arterial Pressure , Dan Roll Over Test Di Puskesmas Pacar Keling Surabaya*. 11–14.

Tahir, M., Badriyah, T., & Syarif, I. (2018). Classification Algorithms of Maternal Risk Detection For Preeclampsia With Hypertension During Pregnancy Using Particle Swarm Optimization. *EMITTER International Journal of Engineering Technology*, *6*(2), 236–253. https://doi.org/10.24003/emitter.v6i2.287

Tarimo, C. S., Bhuyan, S. S., Li, Q., Ren, W., Mahande, M. J., & Wu, J. (2021). Combining resampling strategies and ensemble machine learning methods to enhance prediction of neonates with a low apgar score after induction of labor in Northern Tanzania. *Risk Management and Healthcare Policy*, *14*(September), 3711–3720. https://doi.org/10.2147/RMHP.S331077

Taunk, K. (2019). 2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019. *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, *Iciccs*, 1255–1260.

Thakur, S., Dharavath, R., & Edla, D. R. (2020). Spark and Rule-KNN based scalable machine learning framework for EEG deceit identification. *Biomedical Signal Processing and Control*, *58*, 101886. https://doi.org/10.1016/j.bspc.2020.101886

Thammasiri, D., Delen, D., Meesad, P., & Kasap, N. (2014). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. *Expert Systems with Applications*, *41*(2), 321–330. https://doi.org/10.1016/j.eswa.2013.07.046

Uddin, S., Haque, I., Lu, H., Moni, M. A., & Gide, E. (2022). Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-10358-x

Upadhyay, K., & Kaur, P. (2021). A Review on Data level Approaches to address the Class Imbalance Problem. *International Conference on Challenges in Engineering Science and Technology*, *April*.

Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics*, *85*(June), 189–203. https://doi.org/10.1016/j.jbi.2018.07.014

Valarmathi, R., & Sheela, T. (2021). Heart disease prediction using hyper parameter optimization (HPO) tuning. *Biomedical Signal Processing and Control*, *70*(July), 103033. https://doi.org/10.1016/j.bspc.2021.103033

Vijayashree, J., & Sultana, H. P. (2018). A Machine Learning Framework for Feature Selection in Heart Disease Classification Using Improved Particle Swarm Optimization with Support Vector Machine Classifier. *Programming and Computer Software*, *44*(6), 388–397. https://doi.org/10.1134/S0361768818060129

Wah, Y. B., Rahman, H. A. A., He, H., & Bulgiba, A. (2016). Handling imbalanced dataset using SVM and k-NN approach. *AIP Conference Proceedings*, *1750*(June 2016). https://doi.org/10.1063/1.4954536

Wang, J., & Wang, A. (2023). *Python Tutorial*. 11–41. https://doi.org/10.1007/978-3-031-17646-3_2

Wang, M., Wei, Z., Jia, M., Chen, L., & Ji, H. (2022). Deep learning model for multi-classification of infectious diseases from unstructured electronic medical records. *BMC Medical Informatics and Decision Making*, *22*(1), 1–13. https://doi.org/10.1186/s12911-022-01776-y

Wang, W., & Sun, D. (2021). The improved AdaBoost algorithms for imbalanced data classification. *Information Sciences*, *563*, 358–374. https://doi.org/10.1016/j.ins.2021.03.042

Wang, Y., Wei, Y., Yang, H., Li, J., Zhou, Y., & Wu, Q. (2020). Utilizing imbalanced electronic health records to predict acute kidney injury by ensemble learning and time series model. *BMC Medical Informatics and Decision Making*, *20*(1), 1–13. https://doi.org/10.1186/s12911-020-01245-4

Wang, Z., Jia, P., Xu, X., Wang, B., Zhu, Y., & Li, D. (2021). Sample and feature selecting based ensemble learning for imbalanced problems. *Applied Soft Computing*, *113*, 107884. https://doi.org/10.1016/j.asoc.2021.107884

Weissgerber, T. L., & Mudd, L. M. (2015). Preeclampsia and Diabetes. *Current Diabetes Reports*, *15*(3). https://doi.org/10.1007/s11892-015-0579-4

Wirth, R., & Hipp, J. (2000). CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39. *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, *24959*, 29–39. https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining

Xu, Z., Shen, D., Nie, T., & Kou, Y. (2020a). A hybrid sampling algorithm combining M-SMOTE and ENN based on Random forest for medical imbalanced data. *Journal of Biomedical Informatics*, *107*(May 2019), 103465. https://doi.org/10.1016/j.jbi.2020.103465

Xu, Z., Shen, D., Nie, T., & Kou, Y. (2020b). A hybrid sampling algorithm combining M-SMOTE and ENN based on Random forest for medical imbalanced data. *Journal of Biomedical Informatics*, *107*(May), 103465. https://doi.org/10.1016/j.jbi.2020.103465

Xu, Z., Shen, D., Nie, T., Kou, Y., Yin, N., & Han, X. (2021). A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data. *Information Sciences*, *572*, 574–589. https://doi.org/10.1016/j.ins.2021.02.056

Yildirim, M., & Cinar, A. (2020). Classification of Alzheimer's disease MRI images with CNN based hybrid method. *Ingenierie Des Systemes d'Information*, *25*(4), 413–418. https://doi.org/10.18280/isi.250402

Yu, T., & Zhu, H. (2020). *Hyper-Parameter Optimization: A Review of Algorithms and Applications*. 1–56. http://arxiv.org/abs/2003.05689

Zhang, F., Petersen, M., Johnson, L., Hall, J., & O'bryant, S. E. (2022). Hyperparameter Tuning with High Performance Computing Machine Learning for Imbalanced Alzheimer's Disease Data. *Applied Sciences (Switzerland)*, *12*(13). https://doi.org/10.3390/app12136670

Zhang, J., Han, L., Li, W., Chen, Q., Lei, J., Long, M., Yang, W., Li, W., Zeng, L., & Zeng, S. (2019). Early prediction of preeclampsia and small-for-gestational-age via multi-marker model in Chinese pregnancies: A prospective screening study. *BMC Pregnancy and Childbirth*, *19*(1), 1–10. https://doi.org/10.1186/s12884-019-2455-8

Zhang, X., Chen, Y., Salerno, S., Li, Y., Zhou, L., Zeng, X., & Li, H. (2022). Prediction of severe preeclampsia in machine learning. *Medicine in Novel Technology and Devices*, *15*(July). https://doi.org/10.1016/j.medntd.2022.100158

Zhang, Y., Cao, G., Wang, B., & Li, X. (2019). A novel ensemble method for k-nearest neighbor. *Pattern Recognition*, *85*, 13–25. https://doi.org/10.1016/j.patcog.2018.08.003

Zhang, Z. (2016). Introduction to machine learning: K-nearest neighbors. *Annals of Translational Medicine*, *4*(11). https://doi.org/10.21037/atm.2016.03.37

Zhu, C., & Wang, Z. (2017). Entropy-based matrix learning machine for imbalanced data sets. *Pattern Recognition Letters*, *88*, 72–80. https://doi.org/10.1016/j.patrec.2017.01.014