



## REFERENCES

- Ai, J. *et al.* (2022) “A Systematic Review and Qualitative Assessment of Fraud Detection Methodologies in Health Care,” *North American Actuarial Journal*, 26(1), pp. 1–26. Available at: <https://doi.org/10.1080/10920277.2021.1895843>.
- Bauder, R., Khoshgoftaar, T.M. and Seliya, N. (2017) “A survey on the state of healthcare upcoding fraud analysis and detection,” *Health Services and Outcomes Research Methodology*, 17(1), pp. 31–55. Available at: <https://doi.org/10.1007/S10742-016-0154-8/TABLES/2>.
- Bauder, R.A. and Khoshgoftaar, T.M. (2017) “Medicare fraud detection using machine learning methods,” *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, 2017-December, pp. 858–865. Available at: <https://doi.org/10.1109/ICMLA.2017.00-48>.
- Bush, J. *et al.* (2017) “Medicare fraud, waste and abuse,” *Management Faculty Research* [Preprint]. Available at: [https://mds.marshall.edu/mgmt\\_faculty/176](https://mds.marshall.edu/mgmt_faculty/176) (Accessed: May 4, 2025).
- CatBoost* (no date). Available at: <https://catboost.ai/docs/en/> (Accessed: May 4, 2025).
- Chen, T. and Guestrin, C. (2016) “XGBoost: A scalable tree boosting system,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13-17-August-2016, pp. 785–794. Available at: [https://doi.org/10.1145/2939672.2939785/SUPPL\\_FILE/KDD2016\\_CHE\\_N\\_BOOSTING\\_SYSTEM\\_01-ACM.MP4](https://doi.org/10.1145/2939672.2939785/SUPPL_FILE/KDD2016_CHE_N_BOOSTING_SYSTEM_01-ACM.MP4).
- DecisionTreeClassifier* — *scikit-learn 1.6.1 documentation* (no date). Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html> (Accessed: May 4, 2025).
- Dunford, R. *et al.* (2014) “The Pareto Principle,” *The Plymouth Student Scientist*, 7(1), pp. 140–148. Available at: <https://doi.org/10.24382/swfr-wr17>.
- ExtraTreesClassifier* — *scikit-learn 1.6.1 documentation* (no date). Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.ExtraTreesClassifier.html> (Accessed: May 4, 2025).
- GaussianNB* — *scikit-learn 1.6.1 documentation* (no date). Available at: [https://scikit-learn.org/stable/modules/generated/sklearn.naive\\_bayes.GaussianNB.html](https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.GaussianNB.html) (Accessed: May 4, 2025).
- GradientBoostingClassifier* — *scikit-learn 1.6.1 documentation* (no date). Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html> (Accessed: May 4, 2025).
- Guo, G. *et al.* (2003) “KNN Model-Based Approach in Classification,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial*



- Intelligence and Lecture Notes in Bioinformatics*), 2888, pp. 986–996. Available at: [https://doi.org/10.1007/978-3-540-39964-3\\_62](https://doi.org/10.1007/978-3-540-39964-3_62).
- Ke, G. *et al.* (2017) “LightGBM: A Highly Efficient Gradient Boosting Decision Tree,” *Advances in Neural Information Processing Systems*, 30. Available at: <https://github.com/Microsoft/LightGBM>. (Accessed: May 20, 2025).
- KNeighborsClassifier — scikit-learn 1.6.1 documentation* (no date). Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html> (Accessed: May 4, 2025).
- Kose, I., Gokturk, M. and Kilic, K. (2015) “An interactive machine-learning-based electronic fraud and abuse detection system in healthcare insurance,” *Applied Soft Computing*, 36, pp. 283–299. Available at: <https://doi.org/10.1016/J.ASOC.2015.07.018>.
- LaValley, M.P. (2008) “Logistic regression,” *Circulation*, 117(18), pp. 2395–2399. Available at: <https://doi.org/10.1161/CIRCULATIONAHA.106.682658/ASSET/472996AA-9EBA-4322-A948-192E511AC922/ASSETS/GRAPHIC/19FF1.JPEG>.
- Lavanya, S., S, M.K. and Kumar, M. (2021) “Machine Learning Based Approaches for Healthcare Fraud Detection: A Comparative Analysis.”
- LogisticRegression — scikit-learn 1.6.1 documentation* (no date). Available at: [https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LogisticRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html) (Accessed: May 4, 2025).
- Mahmoud, A.Y. (2024) “Novel efficient feature selection: Classification of medical and immunotherapy treatments utilising Random Forest and Decision Trees,” *Intelligence-Based Medicine*, 10, p. 100151. Available at: <https://doi.org/10.1016/J.IBMED.2024.100151>.
- MLPClassifier — scikit-learn 1.6.1 documentation* (no date). Available at: [https://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.MLPClassifier.html](https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html) (Accessed: May 4, 2025).
- Natekin, A. and Knoll, A. (2013) “Gradient boosting machines, a tutorial,” *Frontiers in Neurorobotics*, 7(DEC), p. 63623. Available at: <https://doi.org/10.3389/FNBOT.2013.00021/BIBTEX>.
- Prokhorenkova, L. *et al.* (2018) “CatBoost: unbiased boosting with categorical features,” *Advances in Neural Information Processing Systems*, 31.
- Prova, N.N.I. (2024) “Healthcare Fraud Detection Using Machine Learning,” *2nd International Conference on Intelligent Cyber Physical Systems and Internet of Things, ICoICI 2024 - Proceedings*, pp. 1119–1123. Available at: <https://doi.org/10.1109/ICOICI62503.2024.10696476>.
- Python API Reference — xgboost 3.1.0-dev documentation* (no date). Available at: [https://xgboost.readthedocs.io/en/latest/python/python\\_api.html](https://xgboost.readthedocs.io/en/latest/python/python_api.html) (Accessed: May 4, 2025).
- Python-package Introduction — LightGBM 4.6.0.99 documentation* (no date). Available at: <https://lightgbm.readthedocs.io/en/latest/Python-Intro.html> (Accessed: May 4, 2025).
- RandomForestClassifier — scikit-learn 1.6.1 documentation* (no date). Available at:



- <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> (Accessed: May 4, 2025).
- de Ville, B. (2013) "Decision trees," *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(6), pp. 448–455. Available at: <https://doi.org/10.1002/WICS.1278>.
- Waghade, S.S. (2018) "A Comprehensive Study of Healthcare Fraud Detection based on Machine Learning," *International Journal of Applied Engineering Research*, 13(6), pp. 4175–4178. Available at: <http://www.ripublication.com> (Accessed: May 1, 2025).
- Webb, G.I. (2017) "Naïve Bayes," *Encyclopedia of Machine Learning and Data Mining*, pp. 895–896. Available at: [https://doi.org/10.1007/978-1-4899-7687-1\\_581](https://doi.org/10.1007/978-1-4899-7687-1_581).