

Fraud Detection and Analysis for Insurance Claim using Machine Learning

Akansh Raj¹, Aditya Raj², Aditya Narayan Sharma³, Saket Kumar⁴, Rohit Kaushik⁵, Shallu Rani⁶

Students, Computer Science and Engineering¹⁻⁵

Supervisor, Computer Science and Engineering⁶

Chandigarh University, Punjab, India

akanshraj616@gmail.com, adityaraj181630@gmail.com, adinarayanofficial@gmail.com,

20BCS7524@cumail.in, rohithawk001@gmail.com, shallu.e15381@cumail.in

Abstract: Insurance Company working as commercial enterprise from last few years have been experiencing fraud cases for all type of claims. Amount claimed by fraudulent is significantly huge that may causes serious problems, hence along with government, different organization also working to detect and reduce such activities. Such frauds occurred in all areas of insurance claim with high severity such as insurance claimed towards auto sector is fraud that widely claimed and prominent type, which can be done by fake accident claim. So, we aim to develop a project that work on insurance claim data set to detect fraud and fake claims amount. The project implement machine learning algorithms to build model to label and classify claim. Also, to study comparative study of all machine learning algorithms used for classification using confusion matrix in term soft accuracy, precision, recall etc. For fraudulent transaction validation, machine learning model is built using PySpark Python Library.

Keywords: Machine Learning Algorithm, PySpark, FraudCase detection, classifications

REFERENCES

- [1] K. Ulaga Priya and S. Pushpa, "A Survey on Fraud Analytics Using Predictive Model in Insurance Claims," Int. J. Pure Appl. Math., vol. 114, no. 7, pp. 755–767, 2017.
- [2] E. B. Belhadji, G. Dionne, and F. Tarkhani, "A Model for the Detection of Insurance Fraud," Geneva Pap. Risk Insur. Issues Pract., vol. 25, no. 4, pp. 517– 538, 2000, doi: 10.1111/1468-0440.00080.
- [3] "Predictive Analysis for Fraud Detection." <https://www.wipro.com/analytics/comparative-analysis-of-machine-learning-techniques-for-%0Adetectin/>.
- [4] F. C. Li, P. K. Wang, and G. E. Wang, "Comparison of the primitive classifiers with extreme learning machine in credit scoring," IEEM 2009 - IEEE Int. Conf. Ind. Eng. Eng. Manag., vol. 2, no. 4, pp. 685– 688, 2009, doi: 10.1109/IEEM.2009.5373241.
- [5] V. Khadse, P. N. Mahalle, and S. V. Biraris, "An Empirical Comparison of Supervised Machine Learning Algorithms for Internet of Things Data," Proc. - 2018 4th Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2018, pp. 1–6, 2018, doi: 10.1109/ICCUBEA.2018.8697476.
- [6] S. Ray, "A Quick Review of Machine Learning Algorithms," Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Perspectives Prospect. Com. 2019, pp. 35–39, 2019, doi: 10.1109/COMITCon.2019.8862451.
- [7] "<https://www.dataschool.io/comparing-supervised-learning-algorithms/>".