

Klasifikasi Penipuan pada Rekening Bank menggunakan Pendekatan Ensemble Learning

Oleh:

Alfiah Maghfiroh,

Yulian Findawati

Progam Studi Teknik Informatika

Universitas Muhammadiyah Sidoarjo

Maret, 2023



Pendahuluan

Rekening merupakan kumpulan nomor yang biasa digunakan untuk semua transaksi di dunia perbankan, mulai dari menabung, tarik tunai, hingga melakukan pengecekan saldo rekening baik secara langsung maupun melalui online menggunakan m-banking. Dalam rekening bank dan proses pembukaanya terdapat upaya tindak kriminal dengan berbagai macam yang dilakukan seseorang maupun kelompok. Sebuah kewajiban bank untuk mencegah terjadinya tindak kriminal penipuan sehingga dapat memberikan kepercayaan pada masyarakat atau nasabah. Dalam upaya pencegahan tindak kriminal penipuan dapat dipecahkan menggunakan teknik data mining yaitu klasifikasi. Tujuan dari klasifikasi yaitu memprediksi label kelas dari suatu objek berdasarkan atribut yang ada. Oleh karena itu penelitian ini dibuat untuk melakukan klasifikasi pada rekening bank supaya dapat dilakukan pencegahan penipuan

Pertanyaan Penelitian (Rumusan Masalah)

1. Bagaimana memprediksi tindak penipuan pada rekening bank menggunakan pendekatan Ensemble Learning?
2. Bagaimana menentukan parameter terbaik menggunakan Randomized searchCV?
3. Bagaimana melakukan uji performa pemodelan prediksi menggunakan confusion matrix ?

Metode

Penelitian ini menggunakan beberapa metode, diantaranya yaitu :

- Randomized Search Cross Validation untuk melakukan tuning parameter
- Minmax Scaller untuk normalisasi data
- Extreme Gradient Boosting dan Random Forest untuk melakukan klasifikasi data
- Confussion Matrix untuk evaluasi model

Hasil

Pengujian pada penelitian ini menggunakan data latih 90% dan data uji 10% yang diterapkan pada dua metode klasifikasi berikut skor yang dihasilkan.

Algoritma	Skor Latih	Skor Uji
Extreme Gradient Boosting (XGBoost)	99,50%	99,59%
Random Forest	99,46%	99,59%

Pembahasan

- Pada tabel tersebut dapat terlihat tahapan penelitian dengan mendapat nilai skor tertinggi didapatkan oleh metode Extreme Gradient Boosting dengan skor Latih 99,50% dan skor uji 99,59%. Sedangkan Random forest mendapat skor latih 99,46% dan skor uji 99,59%.

Temuan Penting Penelitian

- Data yang digunakan pada penelitian adalah data “Bank Account Fraud Dataset suite (NeurIPS 2022) dari Kaggle berjumlah 22029 record.

Manfaat Penelitian

- Penelitian ini diharapkan dapat bermanfaat sebagai acuan dan prediksi atas pencegahan tindak penipuan oleh calon nasabah, sehingga pihak bank dapat mengurangi tindak criminal tersebut, dan sehingga dapat meningkatkan kepercayaan masyarakat pada bank,

Referensi

- [1] R. Chaisari, "Tindak Pidana Penipuan Rekening Bank (Suatu Penelitian Di Wilayah Hukum Polresta Banda Aceh)," Universitas syiah kuala, 2017.
- [2] J. Handoko, K. Kevin, P. Paulus, and Z. Salsabila, "Sistem Deteksi Nomor Telepon dan Rekening Bank Terindikasi Penipuan Berbasis Aplikasi Android dan Web," J. SIFO Mikroskil, vol. 23, no. 2, pp. 183–196, 2022, doi: 10.55601/jsm.v23i2.913.
- [3] A. Prakosa, "Analisis Pengaruh Persepsi Teknologi Dan Persepsi Risiko Terhadap Kepercayaan Pengguna M-Banking," J. Manaj., vol. 9, no. 2, 2019, doi: 10.26460/jm.v9i2.1030.
- [4] L. Septriyana, "Pengambilalihan Kredit Oleh Karyawan Alih Daya (Outsourcing) Pt Bank Mandiri Yang Berakibat Pada Tindak Pidana Penipuan," Indones. Priv. Law Rev., vol. 1, no. 2, pp. 99–106, 2020, doi: 10.25041/iplr.v1i2.2056.
- [5] Suparyanto dan Rosad (2015, "ANALISIS PENERAPAN PRINSIP KEHATI-HATIAN BANK PADA LAYANAN E-BANKING DI PT BANK BNI SYARIAH CABANG MATARAM," Suparyanto dan Rosad (2015, vol. 5, no. 3, pp. 248–253, 2020.
- [6] A. Kurniawan and Y. Yulianingsih, "Pendugaan Fraud Detection pada kartu kredit dengan Machine Learning," Kilat, vol. 10, no. 2, pp. 320–325, 2021, doi: 10.33322/kilat.v10i2.1482.
- [7] M. Febriady, Samsuryadi, and D. P. Rini, "Klasifikasi Transaksi Penipuan Pada Kartu Kredit Menggunakan Metode Resampling Dan Pembelajaran Mesin," J. Media Inform. Budidarma, vol. 6, pp. 1010–1016, 2022, doi: 10.30865/mib.v6i2.3765.
- [8] T. S. Lestari and D. A. N. Sirodj, "Klasifikasi Penipuan Transaksi Kartu Kredit Menggunakan Metode Random Forest," J. Ris. Stat., vol. 1, no. 2, pp. 160–167, 2022, doi: 10.29313/jrs.v1i2.525.
- [9] M. Syukron, R. Santoso, and T. Widiarini, "PERBANDINGAN METODE SMOTE RANDOM FOREST DAN SMOTE XGBOOST UNTUK KLASIFIKASI TINGKAT PENYAKIT HEPATITIS C PADA IMBALANCE CLASS DATA," vol. 9, pp. 227–236, 2020.
- [10] N. N. Pandika Pinata, I. M. Sukarsa, and N. K. Dwi Rusjayanthi, "Prediksi Kecelakaan Lalu Lintas di Bali dengan XGBoost pada Python," J. Ilm. Merpati (Menara Penelit. Akad. Teknol. Informasi), vol. 8, no. 3, p. 188, 2020, doi: 10.24843/jim.2020.v08.i03.p04.
- [11] F. Zamachsari and N. Puspitasari, "Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 5, no. 2, pp. 203–212, 2021, doi: 10.29207/resti.v5i2.2952.
- [12] G. A. Shafila, "IMPLEMENTASI METODE EXTREME GRADIENT BOOSTING (XGBOOST) UNTUK KLASIFIKASI PADA DATA BIOINFORMATIKA (Studi Kasus : Penyakit Ebola , GSE 122692)," Dspace.Uii.Ac.Id, 2020, [Online]. Available: [https://dspace.uui.ac.id/bitstream/handle/123456789/29276/16611022 Gregy Addis Shafila.pdf?sequence=1&isAllowed=y](https://dspace.uui.ac.id/bitstream/handle/123456789/29276/16611022%20Gregy%20Addis%20Shafila.pdf?sequence=1&isAllowed=y).
- [13] J. Paul Mueller and L. Massaron, Algorithm dummies. Canada: john willey&sons,inc., 2017.
- [14] V. YUGESH, "A Complete Guide to Categorical Data Encoding," DEVELOPERS CORNER, 2021. <https://analyticsindiamag.com/a-complete-guide-to-categorical-data-encoding/>.
- [15] O. V. Putra, T. Harmini, and A. Saroji, "Outlier Detection On Graduation Data Of Darussalam Gontor University Using One-Class Support Vector Machine," Procedia Eng. Life Sci., vol. 2, no. 2, pp. 89–92, 2021, doi: 10.21070/pels.v2i0.1139.
- [16] E. Agustin, A. Eviyanti, and N. L. Azizah, "Deteksi Penyakit Epilepsi Melalui Sinyal EEG Menggunakan Metode DWT dan Extreme Gradient Boosting," vol. 7, pp. 117–127, 2023, doi: 10.30865/mib.v7i1.5412.
- [17] S. Shalev-Shwartz and S. Ben-David, Understanding Machine Learning. New york, 2014.
- [18] R. E. Prasetyo, "[Belajar DM] Evaluasi Model Data Mining," 2022. <https://rudyekoprasetya.wordpress.com/2021/04/13/belajar-dm-evaluasi-model-data-mining/>.
- [19] H. Mukhtar et al., "Jurnal Computer Science and Information Technology (CoSciTech) Peramalan Kedatangan Wisatawan ke Suatu Negara Menggunakan Metode Support Vector," vol. 3, no. 3, pp. 274–282, 2022.
- [20] A. Arimuko, A. S. W. Wibawa, and A. Firmansyah, "Analisis Perbandingan Penentuan Hiposentrum Menggunakan Metode Grid Search, Geiger, dan Random Search: Studi Kasus pada Letusan Gunung Sinabung 2017," Diffraction, vol. 1, no. 2, pp. 22–28, 2019, doi: 10.37058/diffraction.v1i2.1290.

