

**IMPLEMENTASI METODE *NAIVE BAYES* DAN *PARTICLE SWARM OPTIMIZATION* (PSO) UNTUK MENDETEKSI KECURANGAN (*FRAUD*)
DALAM KARTU KREDIT**

SKRIPSI

diajukan untuk memenuhi sebagian dari syarat memperoleh gelar
Sarjana Matematika



Oleh:

Hasna Razan Alkhansa

1900512

**PROGRAM STUDI MATEMATIKA
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2023**

LEMBAR PENGESAHAN

HASNA RAZAN ALKHANSA

**IMPLEMENTASI METODE NAIVE BAYES DAN PARTICLE SWARM
OPTIMIZATION (PSO) UNTUK MENDETEKSI KECURANGAN (FRAUD)
DALAM KARTU KREDIT**

disetujui dan disahkan oleh:

Pembimbing I

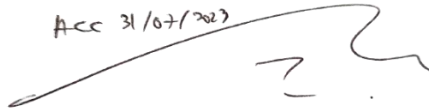
Acc 28/07/2023



Dr. Khusnul Novianingsih, S.Si., M.Si.
NIP.197711282008122001

Pembimbing II

Acc 31/07/2023



Dr. Al Azhary Masta, M.Si.
NIP. 199006102015041001

Mengetahui,

Ketua Program Studi Matematika



Dr. Kartika Yulianti, M.Si.
NIP. 198207282005012001

i

LEMBAR HAK CIPTA

**IMPLEMENTASI METODE *NAIVE BAYES* DAN *PARTICLE SWARM OPTIMIZATION* (PSO) UNTUK MENDETEKSI KECURANGAN (*FRAUD*)
DALAM KARTU KREDIT**

Oleh:

Hasna Razan Alkhansa

1900512

Diajukan untuk memenuhi sebagian dari syarat memperoleh gelar Sarjana Matematika pada Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

© Hasna Razan Alkhansa 2023

Universitas Pendidikan Indonesia

Agustus 2023

Hak cipta dilindungi undang-undang

Skripsi tidak boleh diperbanyak seluruhnya atau sebagian dengan dicetak ulang, difotokopi, atau cara lainnya tanpa izin penulis

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi yang berjudul “Implementasi Metode Naive Bayes dan *Particle Swarm Optimization* (PSO) untuk Mendeteksi Kecurangan (*Fraud*) dalam Kartu Kredit” beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko atau sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya sendiri.

Bandung, Juli 2023

Penulis,



Hasna Razan Alkhansa

NIM 1900512

KATA PENGANTAR

Puji syukur kehadiran Allah SWT karena dengan rahmat, karunia, serta taufik dan hidayah-Nya penulis dapat menyelesaikan skripsi berjudul: “Implementasi Metode *Naive Bayes* dan *Particle Swarm Optimization* (PSO) untuk Mendeteksi Kecurangan (*Fraud*) dalam Kartu Kredit”. Penulisan skripsi ini bertujuan sebagai salah satu syarat yang harus dipenuhi untuk memperoleh gelar Sarjana Matematika di Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan Matematika.

Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada kepada pihak-pihak yang telah memberikan dukungan dan bimbingan selama proses penulisan skripsi ini karena penulisan skripsi ini tidak terlepas dari bantuan dan dukungan berbagai pihak.

Penulis menyadari bahwa penelitian ini masih memiliki keterbatasan dan ruang untuk pengembangan lebih lanjut. Oleh karena itu, penulis berharap penelitian ini dapat menjadi kontribusi yang bermanfaat bagi perkembangan ilmu pengetahuan, khususnya dalam penelitian yang sejenis.

UCAPAN TERIMA KASIH

Penulis menyadari bahwa skripsi ini dapat selesai tepat waktu karena adanya bantuan, bimbingan, arahan, dan motivasi dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih sebanyak-banyaknya kepada:

1. Orang tua dan keluarga tercinta yang telah memberikan doa, motivasi, dan dukungan materi dari awal hingga akhir perkuliahan.
2. Ibu Dr. Khusnul Novianingsih, M.Si., selaku Pembimbing I yang telah memberikan arahan dan bimbingan dalam proses penulisan skripsi ini.
3. Bapak Dr. Al Azhary Masta, M.Si., selaku Pembimbing II yang telah memberikan arahan dan bimbingan dalam proses penulisan skripsi ini.
4. Ibu Dra. Rini Marwati, M.S. selaku dosen Pembimbing Akademik yang telah memberikan arahan serta motivasi selama menjalani perkuliahan ini.
5. Seluruh dosen Program Studi Matematika yang telah memberikan ilmunya dari awal hingga akhir perkuliahan.
6. Teman-teman terdekat, Afif, Sarah, Biba, Zehan, Gita, Bening, dan teman teman RD lainnya, yang telah memberikan semangat dan dukungan setiap harinya.
7. Teman-teman kelas Matematika C 2019, Fanny, Bena, Shafa dan teman teman Purcel lainnya, yang telah kebersamai dan memberikan dukungan selama menjalani proses perkuliahan.
8. Seluruh pihak yang tidak dapat disebutkan satu persatu yang telah membantu penulisan selama studi dan penyusunan skripsi ini.

Semoga Allah SWT senantiasa membalas kebaikan serta ketulusan dari pihak-pihak yang telah membantu penulis.

**IMPLEMENTASI METODE NAIVE BAYES DAN PARTICLE SWARM
OPTIMIZATION (PSO) UNTUK MENDETEKSI KECURANGAN (FRAUD)
DALAM KARTU KREDIT**

ABSTRAK

Kasus kecurangan (*fraud*) terus meningkat seiring dengan meningkatnya penggunaan kartu kredit. Meskipun berbagai metode otorisasi telah digunakan, kecurangan kartu kredit masih sulit untuk dihindari. Penelitian ini bertujuan untuk mengidentifikasi kecurangan (*fraud*) dengan melihat riwayat transaksi kartu kredit menggunakan metode *Naive Bayes*. Proses pengembangan model penelitian dilakukan dengan menggunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Data yang digunakan bersumber dari *Bank Simulation* atau *BankSim*, sebuah simulator pembayaran bank berbasis agen yang menghasilkan data sintetik untuk penelitian deteksi penipuan. Pada tahap awal, dilakukan seleksi fitur menggunakan *Particle Swarm Optimization (PSO)* untuk memilih atribut-atribut yang memiliki pengaruh signifikan terhadap kecurangan. Selanjutnya, metode *Naive Bayes* digunakan untuk melakukan klasifikasi pada data yang telah diproses. Evaluasi dan validasi dilakukan untuk mengukur kinerja model yang dikembangkan. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* dan seleksi fitur dengan PSO memperoleh tingkat keberhasilan deteksi dan prediksi yang tinggi, serta tingkat kesalahan yang relatif rendah. Akurasi model *Naive Bayes* mencapai 92.72%, dengan presisi 93.34%, *recall* 91.40%, dan tingkat kesalahan memprediksi *fraud (fraud prediction error)* 6.54%.

Kata Kunci: Kecurangan, Kartu Kredit, *Naive Bayes*, *Particle Swarm Optimization*, Seleksi Fitur.

IMPLEMENTATION OF NAIVE BAYES AND PARTICLE SWARM OPTIMIZATION (PSO) METHODS TO DETECT FRAUD IN CREDIT CARDS

ABSTRACT

Fraud cases continue to increase along with the increasing use of credit cards. Although various authorization methods have been used, credit card fraud is still difficult to avoid. This research aims to identify fraud according to credit card transaction history using the Naive Bayes method. The development of model was implemented using the CRISP-DM (Cross-Industry Standard Process for Data Mining). The data of this research is from Bank Simulation or BankSim, an agent-based bank payment simulator that generates synthetic data for fraud detection research. In the initial stage, feature selection is performed using Particle Swarm Optimization (PSO) to select attributes that significantly influence fraud. Next, the Naive Bayes method is used to classify the processed data. Evaluation and validation were conducted to measure the performance of the developed model. The computational results showed that the Naive Bayes method and feature selection with PSO is successfully implemented with high detection and prediction success rates and relatively low error rates. The accuracy of the Naive Bayes model reached 92.72%, with 93.34% precision, 91.40% recall, and a fraud prediction error rate of 6.54%.

Keywords: *Fraud, Credit Card, Naive Bayes, Particle Swarm Optimization, Feature Selection.*

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR HAK CIPTA	ii
LEMBAR PERNYATAAN	iii
KATA PENGANTAR	iv
UCAPAN TERIMA KASIH.....	v
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
DAFTAR LAMPIRAN.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
BAB II LANDASAN TEORI	5
2.1 <i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i>	5
2.2 Optimasi Metaheuristik	6
2.3 <i>Particle Swarm Optimization (PSO)</i>	7
2.4 <i>Naive Bayes</i>	9
BAB III METODOLOGI PENELITIAN.....	11
3.1 Deskripsi Masalah	11
3.2 Data Penelitian	12
3.3 Pengolahan Data.....	12
3.4 Algoritma PSO untuk Pembobotan Atribut.....	15
3.5 <i>Naive Bayes Classifier (NBC)</i>	17
3.6 Validasi dan Evaluasi	19
3.7 Contoh Kasus	20
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Analisis Data Penelitian	28

4.2	<i>Data Preparation</i>	32
4.3	Pemodelan dengan <i>Naive Bayes</i>	37
4.4	Validasi dan Evaluasi	38
BAB V KESIMPULAN.....		42
5.1	Kesimpulan.....	42
5.2	Saran.....	42
DAFTAR PUSTAKA		44
LAMPIRAN.....		46

DAFTAR GAMBAR

Gambar 2.1 <i>The CRISP-DM Process Model of Data Mining</i>	5
Gambar 3.1 <i>Confusion Matrix</i>	19
Gambar 4.1 Distribusi <i>Fraud</i>	30
Gambar 4.2 Distribusi <i>Age</i>	30
Gambar 4.3 Distribusi <i>Gender</i>	31
Gambar 4.4 Jumlah Transaksi <i>Fraud</i> dan Tidak <i>Fraud</i>	32
Gambar 4.5 Proses <i>Data Cleaning</i>	33
Gambar 4.6 Pengecekan <i>Missing Value</i>	33
Gambar 4.7 Proses Transformasi <i>Dataset</i>	34
Gambar 4.8 Korelasi Antar Fitur dengan Kolom <i>Fraud</i>	35
Gambar 4.9 Hasil PSO.	35
Gambar 4.10 Data Sebelum SMOTE.....	36
Gambar 4.11 Data Setelah SMOTE.	36
Gambar 4.12 Algoritma <i>Naive Bayes</i>	37
Gambar 4.13 Evaluasi Model dengan <i>10-fold Cross Validation</i>	39
Gambar 4.14 Hasil <i>Confusion Matrix</i>	40
Gambar 4.15 Hasil Klasifikasi NBC.....	40

DAFTAR TABEL

Tabel 3.1 Atribut-Atribut <i>Dataset</i>	12
Tabel 3.2 <i>Dataset</i> Transaksi.....	21
Tabel 3.3 Matriks A	21
Tabel 3.4 Matriks B	22
Tabel 3.5 Nilai Sigmoid.....	22
Tabel 3.6 Hasil Perbandingan A dengan Nilai Sigmoid.....	22
Tabel 3.7 Hasil <i>Pbest</i>	23
Tabel 3.8 Kecepatan Partikel Iterasi ke Dua.....	24
Tabel 3.9 Posisi Partikel Iterasi ke Dua.....	24
Tabel 3.10 <i>Dataset</i> Transaksi Setelah Diseleksi.....	25
Tabel 3.11 Probabilitas Setiap Atribut.....	26
Tabel 4.1 <i>Dataset</i> Penelitian.....	28
Tabel 4.2 Tipe Data dari Setiap Kolom dari Data Penelitian.....	29
Tabel 4.3 Distribusi Rata-rata <i>Fraud</i> Berdasarkan <i>Category</i>	31
Tabel 4.4 Tampilan <i>Dataset</i> Setelah <i>Cleaning</i>	34
Tabel 4.5 Hasil Model <i>Naive Bayes</i>	38

DAFTAR LAMPIRAN

Lampiran 1: Riwayat Hidup Penulis.....	46
Lampiran 2: <i>Coding Python</i> dalam Proses Seleksi Fitur.	47
Lampiran 3: <i>Coding Python</i> dalam Proses Penyeimbangan Data.....	49
Lampiran 4: <i>Coding Python</i> dalam Proses Model <i>Naive Bayes</i>	51
Lampiran 5: <i>Coding Python</i> dalam Proses Validasi dan Evaluasi.....	52

DAFTAR PUSTAKA

- Abualigah, L. M., & Khader, A. T. (2017). Unsupervised text feature selection technique based on hybrid particle swarm optimization algorithm with genetic operators for the text clustering. *Journal of Supercomputing*, 73 (11), 4773–4795. doi:<https://doi.org/10.1007/s11227-017-2046-2>
- Bhowmik, R. (2008). Data Mining Techniques in Fraud Detection. *Digital Forensics Security and Law*, 3 (2), 35.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. The CRISP-DM consortium.
- Cholissodin, I., & Riyandani, E. (2016). *Swarm Intelligence (Teory & Case Study)*. Malang: Fakultas Ilmu Komputer, Universitas Brawijaya.
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. Diane Cerra.
- Hussain, K., Salleh, M. N., Cheng, S., & Shi, Y. (2018). Metaheuristic research: a comprehensive survey. *Artificial Intelligence Review*, 52, 1-43.
- Kamaruddin, S., & Ravi, V. (2016). Credit Card Fraud Detection using Big Data Analytics: Use of PSOANN based One-Class Classification. *Proceedings of the International Conference on Informatics and Analytics*(33), 1-8. doi:<https://dl.acm.org/doi/10.1145/2980258.2980319>
- Li, J., Ding, L., & Li, B. (2014). A Novel Naive Bayes Classification Algorithm Based on Particle Swarm Optimization. *The Open Automation and Control Systems Journal*, 6, 747-753.
- Lopez-Rojas, E., & Axelsson, S. (2014). Banksim: A bank payments simulator for fraud detection research. *26th European Modeling and Simulation Symposium* (pp. 144-152). Bordeaux, France: Researchgate.
- Muhamad, H., Prasajo, C. A., Sugianto, N. A., Surtiningsih, L., & Cholissodin, I. (2017, September). Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 4, 180-184.
- Nugroho, K. S. (2019, November 13). *Confusion Matrix untuk Evaluasi Model pada Supervised Learning*. Dipetik 2022, dari Medium: <https://ksnugroho.medium.com/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f>
- Prabowo, A. D., & Muljono, M. (2018). Prediksi Nasabah Yang Berpotensi Membuka Simpanan Deposito Menggunakan Naive Bayes Berbasis Particle Swarm Optimization. *Techno.COM*, 17 (2), 208-219.

- Prasetyo, E. (2012). *Data Mining : Konsep dan Aplikasi menggunakan MATLAB* (1st ed.). Yogyakarta: CV Andi Offset.
- Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit Card Fraud Detection Using AdaBoost and Majority Voting. *IEEE Access*, 6, 14277–14284. doi:<http://dx.doi.org/10.1109/ACCESS.2018.2806420>
- Rojas, E. L. (2017, Juli). *Synthetic data from a financial payment system*. Dipetik Mei 24, 2023, dari Kaggle.com: <https://www.kaggle.com/datasets/ealaxi/banksim1?datasetId=1565&sortBy=voteCount>
- Shi, Y., & Eberhart, R. (1998). Modified Particle Swarm Optimizer. *IEEE International Conference on Evolutionary Computation*, 69-73.
- Siringoringo, R. (2018). Klasifikasi Data Tidak Seimbang Menggunakan Algoritma SMOTE dan k-Nearest Neighbor. *Journal Information System Development (ISD)*, 3 (1), 44-49.
- Yang, X. (2010). *Nature-Inspired Metaheuristic Algorithms Second Edition*. United Kingdom: Luniver Press.