

PENGEMBANGAN MODEL KLASIFIKASI PENIPUAN KARTU KREDIT DALAM TRANSAKSI ONLINE MENGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

Haikal Fikri ^{#1}, Sri Hartanto ^{#2}, Herry Wahyono ^{#3}

*Program Studi Teknik Informatika, Universitas Krisnadwipayana
Kampus UNKRIS Jatiwaringin, Pondok Gede 13077*

¹ hfartificial23@gmail.com

² srihartanto@unkris.ac.id

³ wahyonos2000@unkris.ac.id

Abstract

Penipuan kartu kredit dalam transaksi online merupakan ancaman serius yang menyebabkan kerugian finansial bagi konsumen dan lembaga keuangan. Penelitian ini bertujuan mengembangkan model deteksi penipuan kartu kredit menggunakan algoritma *Support Vector Machine* (SVM). Prosesnya meliputi *preprocessing* data seperti penanganan *missing values*, *encoding* fitur kategorikal, dan normalisasi fitur numerik. Dataset dibagi menjadi data pelatihan dan pengujian untuk validasi model. Evaluasi model dilakukan menggunakan akurasi, *precision*, recall, dan *F1-score*. Hasil menunjukkan model SVM mencapai akurasi, *precision*, recall, dan *F1-score* sebesar 100%. Validasi silang memastikan konsistensi, dan optimasi hyperparameter dengan GridSearchCV meningkatkan kinerja model. Studi ini menyimpulkan bahwa model SVM yang dihasilkan andal untuk deteksi penipuan kartu kredit.

Keywords: Penipuan kartu kredit, Support Vector Machine, deteksi penipuan, transaksi online, hyperparameter

I. INTRODUCTION

Dalam era digital saat ini, transaksi online menjadi bagian penting dari kehidupan sehari-hari, dan kartu kredit adalah salah satu metode pembayaran yang paling banyak digunakan. Namun, meningkatnya transaksi *online* juga disertai dengan peningkatan risiko penipuan, yang mencakup penipuan identitas, transaksi, dan keamanan, yang dapat menyebabkan kerugian finansial dan reputasi bagi pihak yang terlibat. Kartu kredit,

sebagai alat pembayaran dari bank atau perusahaan untuk transaksi barang atau jasa, menawarkan berbagai kemudahan dan manfaat menarik yang membuat banyak orang memilih untuk menggunakannya [1].

Data dari Bank Indonesia menunjukkan bahwa jumlah kartu kredit di Indonesia meningkat secara signifikan dalam dekade terakhir, mencapai lebih dari 16 juta kartu pada tahun 2022. *Volume* transaksi pun mengalami peningkatan yang mencolok, yang mengindikasikan ketergantungan yang terus bertambah pada kartu kredit [2]. Tingginya angka penipuan dalam transaksi ini menunjukkan kebutuhan akan sistem deteksi penipuan otomatis, yang menggunakan teknik pembelajaran mesin untuk mendeteksi transaksi yang mencurigakan secara real-time [3].

Model *machine learning* seperti *Support Vector Machine* (SVM) sangat relevan untuk mengatasi tantangan ini karena kemampuan klasifikasinya, baik linier maupun *non-linier*, serta kemampuannya dalam meminimalkan kesalahan klasifikasi. Namun, dalam pengembangannya, tantangan ketidakseimbangan data antara transaksi *fraud* dan *non-fraud* menjadi salah satu isu utama yang dapat memengaruhi akurasi dan konsistensi model prediksi [3]. Penelitian ini bertujuan untuk mengembangkan model prediksi penipuan kartu kredit yang efektif dan akurat menggunakan algoritma SVM untuk mendeteksi transaksi tidak sah dan mencegah kerugian finansial serta reputasi bagi pihak yang terlibat.

II. LITERATURE REVIEW

Analisis sentimen merupakan teknik yang digunakan untuk menentukan sikap atau opini individu terhadap suatu entitas, baik itu produk, layanan, atau isu tertentu. Salah satu metode yang banyak digunakan dalam analisis sentimen adalah *Support Vector Machine* (SVM). SVM merupakan algoritma pembelajaran mesin yang diperkenalkan oleh Cortes dan Vapnik pada tahun 1995. Dalam analisis sentimen, SVM berfungsi untuk memisahkan data ke dalam dua kategori, seperti positif dan negatif, dengan menemukan *hyperplane* optimal yang memisahkan kedua kelas tersebut [2].

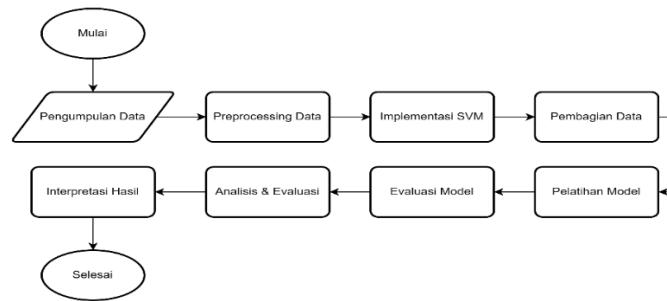
Salah satu keunggulan SVM adalah kemampuannya untuk mengatasi data yang tidak seimbang. Hal ini sangat penting, mengingat data sentimen sering kali memiliki proporsi kelas yang tidak seimbang antara sentimen positif dan negatif. Dalam penelitian oleh Joachims (1998), SVM menunjukkan performa yang lebih baik dibandingkan algoritma lain seperti Naïve Bayes dan Decision Trees dalam klasifikasi teks [4]. Penelitian ini juga menyoroti efektivitas SVM dalam menangani dimensi fitur yang tinggi, yang merupakan karakteristik umum dalam analisis sentimen dari data teks.

Lebih lanjut, penerapan SVM dalam analisis sentimen telah dilakukan di berbagai domain, termasuk media sosial dan ulasan produk. Fitri dan Putri (2022) mengimplementasikan SVM untuk menganalisis sentimen pengguna terhadap aplikasi Google Meet di Twitter. Hasil penelitian menunjukkan bahwa SVM dapat mencapai akurasi hingga 94%, menandakan kemampuannya dalam menangkap opini publik secara efektif [5]. Selain itu, penelitian oleh Auliya Rahman Isnain et al. (2021) yang membandingkan SVM dengan beberapa algoritma klasifikasi lainnya dalam konteks analisis kebijakan publik juga mencatat bahwa SVM memberikan hasil yang superior dalam hal presisi dan recall, terutama dalam kondisi dataset yang tidak seimbang [6].

Secara keseluruhan, SVM adalah salah satu algoritma yang paling banyak digunakan dalam analisis sentimen, dengan berbagai aplikasi yang menunjukkan hasil yang menjanjikan. Penelitian lebih lanjut diharapkan dapat mengeksplorasi kombinasi SVM dengan teknik pembelajaran mesin lainnya dan pengolahan bahasa alami untuk meningkatkan akurasi dan efisiensi dalam analisis sentimen.

III. RESEARCH METHOD

Tahapan penelitian yang dilakukan adalah sebagai berikut:



Gambar 1. Tahapan Penelitian

Berdasarkan Gambar di atas, terdapat 8 tahapan yang penulis lakukan dalam penelitian ini, yaitu:

1. Pengumpulan Data

Proses ini melibatkan pengumpulan data yang relevan untuk penelitian. Data dapat diperoleh dari sumber primer seperti survei atau wawancara, atau dari sumber sekunder seperti *database* atau media sosial.

2. *Preprocessing* Data

Tahap ini mencakup langkah-langkah pembersihan dan persiapan data agar siap untuk analisis. Ini termasuk penghapusan data duplikat, penanganan nilai hilang, normalisasi, penghapusan *stopwords*, tokenisasi, dan stemming, serta konversi data ke format yang sesuai.

3. Implementasi SVM (*Support Vector Machine*)

Pada tahap ini, algoritma SVM diterapkan untuk membangun model klasifikasi. SVM bekerja dengan mencari *hyperplane* yang memisahkan data dengan cara yang optimal, sehingga meminimalkan kesalahan klasifikasi.

4. Pembagian Data

Data yang telah diproses dibagi menjadi dua set: *set* pelatihan dan *set* pengujian. *Set* pelatihan digunakan untuk melatih model, sedangkan *set* pengujian digunakan untuk mengevaluasi kinerja model pada data yang tidak terlihat sebelumnya.

5. Pelatihan Model

Model SVM dilatih menggunakan *set* pelatihan. Selama pelatihan, model belajar dari pola dalam data untuk memprediksi kelas atau kategori dari data yang akan datang berdasarkan fitur yang ada.

6. Evaluasi Model

Model yang telah dilatih dievaluasi menggunakan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Tahap ini bertujuan untuk menilai seberapa baik model dapat memprediksi kelas pada *set* pengujian.

7. Analisis & Evaluasi

Setelah evaluasi, analisis lebih lanjut dilakukan untuk memahami hasil yang diperoleh. Ini mencakup analisis kesalahan, mengidentifikasi *area* yang perlu diperbaiki, serta membandingkan model dengan algoritma lain jika diperlukan.

8. Interpretasi Hasil

Tahap terakhir adalah menginterpretasikan hasil analisis dan evaluasi. Penulis akan menjelaskan implikasi hasil penelitian, memberikan wawasan terhadap data, serta menjawab pertanyaan penelitian berdasarkan temuan yang diperoleh.

IV. RESULTS AND DISCUSSION

Hasil penelitian menunjukkan bahwa model SVM memiliki kinerja yang baik dalam mendeteksi penipuan kartu kredit, dengan tingkat *accuracy*=100%, *precision*=100%, *recall* =100%, dan *f1-score*=100%. Teknik validasi silang digunakan untuk memastikan konsistensi model, dan peningkatan model dilakukan dengan optimasi *hyperparameter* menggunakan GridSearchCV. Validasi akhir menunjukkan bahwa model yang dihasilkan memiliki kinerja yang optimal dan dapat diandalkan untuk digunakan pada data nyata.

Preprocessing Data

Proses ini mencakup berbagai langkah yang bertujuan untuk mempersiapkan data mentah menjadi data yang sesuai untuk digunakan dalam model *machine learning*, adapun hasilnya sebagai berikut:

```
      Client_Num Card_Category ... Interest_Earned
Delinquent_Acc
0      700082083      Blue ...      4393.21
0
1      700083283      Blue ...      69.44
0
2      700084558      Blue ...      202.58
0
3      700085458      Blue ...      236.40
0
4      700086958      Blue ...      1004.87
1

[5 rows x 18 columns]
```

Gambar 2. Preprocessing Data

Normalisasi Data

Normalisasi data penting untuk memastikan bahwa semua fitur berada dalam skala yang sama. Fitur-fitur numerik seperti *'amt'*, *'lat'*, *'long'*, *'city_pop'*, *'merch_lat'*, dan *'merch_long'* dinormalisasi menggunakan StandardScaler. Adapun hasilnya sebagai berikut:

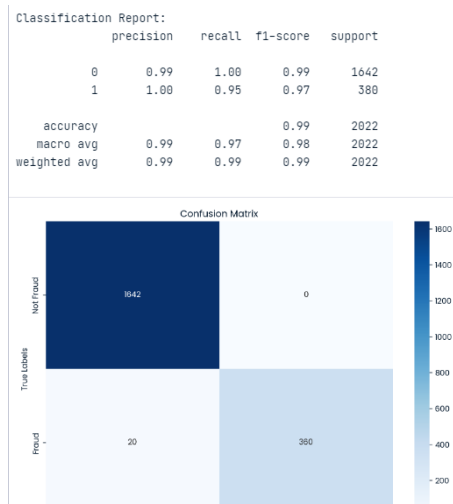
```
      Card_Category_Blue Card_Category_Gold ... Interest_Earned
Delinquent_Acc
0      1.0      0.0 ...      4393.21
0.0
1      1.0      0.0 ...      69.44
0.0
2      1.0      0.0 ...      202.58
0.0
3      1.0      0.0 ...      236.40
0.0
4      1.0      0.0 ...      1004.87
1.0

[5 rows x 18 columns]
```

Gambar 3. Normalisasi Data

Pelatihan Model SVM

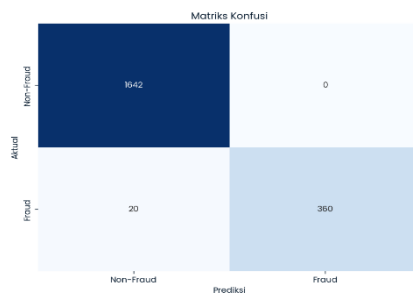
Model SVM dipilih karena kemampuannya dalam memisahkan data dengan margin yang maksimal, yang sangat berguna dalam mendeteksi anomali seperti penipuan kartu kredit. Adapun hasilnya sebagai berikut:



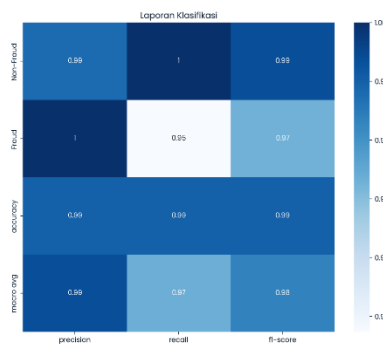
Gambar 4. Pelatihan Model SVM

Pengujian Sistem

Kinerja model dievaluasi menggunakan berbagai metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Matriks konfusi dan laporan klasifikasi digunakan untuk memberikan gambaran yang lebih detail mengenai kinerja model, adapun hasilnya sebagai berikut:



Gambar 5. Pengujian Sistem Matriks Konfusi



Gambar 6. Pengujian Sistem Laporan Klasifikasi

Peningkatan Model

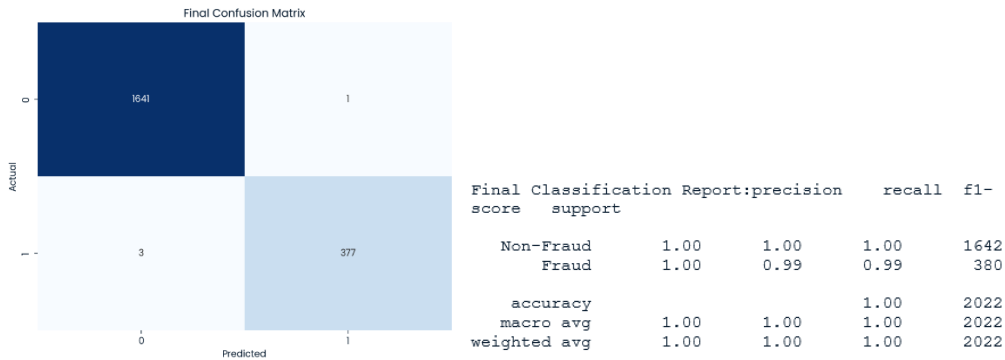
Model ditingkatkan dengan menyesuaikan parameter atau mencoba algoritma lain untuk melakukan pencarian *hyperparameter* menggunakan GridSearchCV untuk meningkatkan model SVM. Adapun hasilnya sebagai berikut:

```
Best Parameters: {'C': 100, 'gamma': 1, 'kernel':  
'linear'}  
  
Classification Report for Best Model:  
              precision    recall  f1-score   support  
  
   Non-Fraud       1.00      1.00      1.00     1642  
     Fraud        1.00      0.99      0.99      380  
  
 accuracy              1.00      2022  
 macro avg           1.00      1.00      1.00     2022  
 weighted avg        1.00      1.00      1.00     2022
```

Gambar 7. Peningkatan Model

Validasi Akhir

Validasi akhir untuk memastikan bahwa model memiliki kinerja yang optimal dan dapat diandalkan untuk digunakan pada data nyata. Validasi akhir ini mencakup evaluasi kinerja model terbaik pada data pengujian. Adapun hasilnya sebagai berikut:



Gambar 8. Matriks Konfusi Validasi Akhir

V. Conclusion

Penelitian ini menunjukkan bahwa algoritma *Support Vector Machine* (SVM) efektif dalam mendeteksi penipuan kartu kredit dalam transaksi *online*, dengan mencapai akurasi 100% pada data pengujian. Proses *preprocessing* data, yang meliputi penanganan nilai hilang, *encoding* fitur kategorikal, dan normalisasi data, berperan penting dalam meningkatkan kinerja model. Validasi silang mengonfirmasi bahwa model SVM tidak mengalami *overfitting* dan menunjukkan kinerja konsisten di berbagai *subset* data. Melalui teknik GridSearchCV, kombinasi parameter terbaik berhasil ditemukan, sehingga memperkuat kemampuan deteksi penipuan. Hasil ini menegaskan bahwa sistem berbasis SVM dapat diandalkan untuk implementasi pencegahan penipuan dalam transaksi kartu kredit secara online.

ACKNOWLEDGMENT

Terimakasih kepada Universitas Krisnadwipayana atas dukungan yang diberikan penelitian ini.

REFERENCES

- [1] Armiani, R., & Agustini, E. P. (2022). Analisa Fraud Pada Transaksi Kartu Kredit Menggunakan Algoritma Random Forest. In *Jurnal Teknologi Informasi dan Terapan (J-TIT)* (Vol. 9, Issue 2). <https://doi.org/10.25047/jtit.v9i2.297>
- [2] Ningsih, P. T. S., Gusvarizon, M., & Hermawan, R. (2022). Analisis Sistem Pendeteksi Penipuan Transaksi Kartu Kredit dengan Algoritma Machine Learning. *Jurnal Teknologi Informatika Dan Komputer*, 8(2), 386–401. <https://doi.org/10.37012/jtik.v8i2.1306>
- [3] Dharmana, I. W., Gunadi, I. G. A., & Dewi, L. J. E. (2024). Deteksi Transaksi *Fraud* Kartu Kredit Menggunakan *Oversampling* ADASYN dan Seleksi Fitur SVM-RFECV. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 125–134. <https://doi.org/10.25126/jtiik.20241117640>
- [4] Romadoni, F., Umidah, Y., & Sari, B. N. (2020). Text Mining Untuk Analisis Sentimen Pelanggan Terhadap Layanan Uang Elektronik Menggunakan Algoritma Support Vector Machine. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 9(2), 247–253. <https://doi.org/10.32736/sisfokom.v9i2.903>
- [5] Angraina, D., & Putri, A. (2022). Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine. *Jurnal CoSciTech (Computer Science and Information Technology)*, 3(3), 472–478. <https://doi.org/10.37859/coscitech.v3i3.4260>
- [6] Rahman Isnain, A., Indra Sakti, A., Alita, D., & Satya Marga, N. (2021). SENTIMEN ANALISIS PUBLIK TERHADAP KEBIJAKAN LOCKDOWN PEMERINTAH JAKARTA MENGGUNAKAN ALGORITMA SVM. *JDMSI*, 2(1), 31–37. <https://t.co/NfhmfMjtXw>