ISSN: 2477-5126
e-ISSN: 2548-9356

Klasifikasi Nasabah Potensial menggunakan Algoritma Ensemble Least Square Support Vector Machine dengan AdaBoost

Benny Leonard Enrico Panggabean[1], Firman Aziz[2*)]
[1,2]Jurusan Ilmu Komputer, Fakultas MIPA, Universitas Pancasakti, Makassar
[1,2]Jln. Andi Mangerangi No. 73, Kota Makassar, 90121, Indonesia
email: [1]blep@unpacti.ac.id, [2]firman.aziz@unpacti.ac.id

Abstract — In the era of business and economics that are interconnected with each other and competition between companies in seeking market share so that there will be an increase, especially in the number of customers, especially deposit customers, financial institutions and other companies are increasingly realizing the importance of understanding and identifying potential customers correctly to get potential customers. customers subscribe to deposits. Potential customer classification is a strategic approach that allows financial institutions to identify potential customers who have the potential to subscribe to deposits. With a deeper understanding of the characteristics and needs of potential customers, financial institutions can direct marketing resources more effectively, increase marketing efforts, and increase the conversion of potential customers to active customers. The aim of this research is to develop and test the Ensemble Least Square Support Vector Machine model with AdaBoost in classifying potential customers which can increase accuracy in identifying potential customers who have the potential to subscribe to deposits. The research results showed that this method achieved an accuracy of 95.15%, a sensitivity of 92.93%, and a specificity of 97.61%. In comparison with single Support Vector Machine and Least Squares Support Vector Machine models, the Ensemble Least Squares Support Vector Machine outperforms both in terms of accuracy.

Abstrak – Dalam era bisnis dan ekonomi yang saling terhubung antara satu sama lain dan persaingan antara perusahaan-perusahaan dalam mencari pangsa pasar agar terjadi peningkatan terutama jumlah nasabah khususnya pelanggan deposito, lembaga keuangan dan perusahaan lainnya semakin menyadari pentingnya memahami dan mengidentifikasi nasabah potensial secara tepat untuk mendapatkan calon nasabah berlangganan deposito. Klasifikasi nasabah potensial adalah suatu pendekatan strategis yang memungkinkan lembaga keuangan untuk mengidentifikasi calon nasabah yang memiliki potensi untuk berlangganan deposito. Dengan pemahaman yang lebih mendalam tentang karakteristik dan kebutuhan nasabah potensial, lembaga keuangan dapat mengarahkan sumber daya bagian marketing dengan lebih efektif, meningkatkan upaya pemasaran, dan meningkatkan konversi nasabah potensial menjadi nasabah aktif. Tujuan penelitian ini adalah untuk mengembangkan dan menguji model Ensemble Least Square Support Vector Machine dengan AdaBoost dalam mengklasifikasi nasabah potensial yang dapat meningkatkan akurasi dalam mengidentifikasi calon nasabah yang berpotensi untuk berlangganan deposito. Hasil penelitian menunjukkan bahwa metode ini mencapai akurasi sebesar 95.15%, sensitivitas sebesar 92.93%, dan spesifisitas sebesar 97.61%. Dalam perbandingan dengan model Support Vector Machine dan Least Square Support Vector Machine tunggal, Ensemble Least Squares Support Vector Machine mengungguli keduanya dalam hal akurasi.

*) **penulis korespondensi**: Firman Aziz
Email: firman.aziz@unpacti.ac.id

Kata Kunci **:SVM,LS-SVM,Ensemble,AdaBoost,Bank Marketing.**

I. PENDAHULUAN

Dalam era globalisasi dan persaingan bisnis yang semakin ketat, lembaga keuangan dan perusahaan lainnya semakin menyadari pentingnya memahami dan mengidentifikasi nasabah potensial dengan tepat. Klasifikasi nasabah potensial adalah suatu pendekatan strategis yang memungkinkan lembaga keuangan untuk mengidentifikasi calon nasabah yang memiliki potensi untuk menggunakan produk atau layanan mereka. Dengan pemahaman yang lebih mendalam tentang karakteristik dan kebutuhan nasabah potensial, lembaga keuangan dapat mengarahkan sumber daya dengan lebih efektif, meningkatkan upaya pemasaran, dan meningkatkan konversi nasabah potensial menjadi nasabah aktif. Ini berarti lembaga keuangan dapat menyesuaikan produk dan layanan mereka dengan lebih baik sesuai dengan preferensi, tujuan, dan tantangan finansial yang dihadapi oleh nasabah potensial, menciptakan pengalaman yang lebih relevan dan memuaskan.

Perkembangan teknologi informasi dan akses internet telah mengubah cara konsumen berinteraksi dengan perusahaan, termasuk lembaga keuangan. Data dan informasi tentang perilaku konsumen kini lebih tersedia dan terukur daripada sebelumnya. Dalam konteks ini, klasifikasi nasabah potensial menjadi semakin relevan, karena lembaga keuangan dapat menggunakan data ini untuk mengidentifikasi prospek yang paling berpotensi untuk menghasilkan pertumbuhan bisnis. Meskipun klasifikasi nasabah potensial sangat penting tetapi penelitian yang komprehensif tentang pendekatan klasifikasi ini masih terbatas. Banyak lembaga keuangan hanya mengandalkan data historis atau insting pasar dalam mengenali nasabah potensial, tanpa memanfaatkan potensi data besar dan teknologi analisis yang lebih canggih. Beberapa penelitian terkait klasifikasi nasabah potensial telah banyak dilakukan tetapi fokus penelitian ini adalah pengembangan metode machine learning menggunakan teknik ensemble.

Dalam penelitian ini, kami akan menggali lebih dalam tentang metode Ensemble Least Square Support Vector Machine dalam klasifikasi nasabah potensial yang lebih akurat dan efisien, dengan mempertimbangkan faktor-faktor

seperti perilaku konsumen, preferensi, karakteristik demografis, dan pola transaksi. Tujuan utama dari penelitian ini adalah untuk mengembangkan dan menguji model Ensemble Least Square Support Vector Machine dengan AdaBoost dalam mengklasifikasi nasabah potensial untuk berlangganan deposito yang dapat meningkatkan akurasi dalam mengidentifikasi calon nasabah yang paling berpotensi. Dalam mencapai tujuan ini, kami akan mempertimbangkan berbagai variabel dan data yang relevan untuk menciptakan model prediktif yang tepat.

II. PENELITIAN YANG TERKAIT

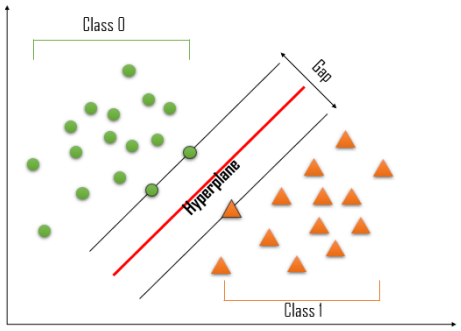
Penelitian [1]–[3] mengadopsi teknik data mining melalui SPSS Modeler untuk memprediksi perilaku langganan deposito berjangka pelanggan dan memahami fitur-fitur pelanggan untuk meningkatkan efektivitas dan keakuratan pemasaran bank. Penelitian ini terbatas hanya pada penarikan kesimpulan tanpa memiliki pengukuran kinerja dalam teknik data mining. Kemudian, penelitian [2] menganalisis dua metode pemilihan fitur yaitu information gain dan Chi-square pada dataset bank direct marketing untuk mengetahui fitur-fitur yang berpengaruh. Hasil menunjukkan metode pemilihan fitur Information Gain dan Chi-square sangat dekat, meskipun keduanya berbeda untuk lima fitur teratas. Penelitian [3] mengusulkan metode untuk menganalisis informasi asimetris menggunakan algoritma SMOTE dan Rotation Forest (PCA) -J48 untuk menyelesaikan klasifikasi set data yang tidak seimbang. dimana algoritma SMOTE digunakan untuk memodifikasi data dan meningkatkan akurasi prediksi. Hasil menunjukkan algoritma SMOTE secara efektif menyelesaikan ketidaksetaraan data dan Rotation Forest (PCA) -J48 dapat mencapai nilai akurasi dan spesifisitas tertinggi. Namun, Rotation Forest (PCA) -J48 memiliki sensitivitas yang tinggi. Penelitian [4] memprediksi respon pelanggan terhadap bank direct marketing dengan menerapkan empat pengklasifikasi yaitu, Multilayer Perceptron Neural Network (MLPNN), Decision Tree (C4.5), Logistic Regression and Random Forest (RF). Hasil menunjukkan bahwa Pengelompokan Random Forest (RF) paling produktif dalam hal kemampuan prediksi dengan akurasi 87%. Kemudian, [5] mengusulkan algoritma klasifikasi Logistic Regression, Decision Tree, Naïve Bayes and the Random Forest ensemble yang digunakan untuk memodelkan dataset. Algoritma tersebut diterapkan pada data bank seimbang mengikuti Cross Industry Standard for Data Mining (CRISP-DM) dan ten-fold cross-validation method. Hasil menunjukkan bahwa duration, poutcome, contact, month dan housing adalah fitur yang paling penting yang berkontribusi pada keberhasilan dalam pemasaran pelanggan bank untuk berlangganan deposit. Penelitian [6] mengusulkan algoritma deep learning untuk mengklasifikasikan data perbankan. Hasil penelitian menunjukkan bahwa metode yang di usulkan memiliki kinerja akurasi sebesar 80%. Tetapi penelitian ini masih perlu pendataan yang lebih kompleks agar prediksi mampu meyakinkan kinerja deep learning berjalan dengan baik. Penelitian [7] mengusulkan metode extraksi fitur t-SNE (t-distributed stochastic neighbor embedding) yang memiliki karakteristik nonlinier dimensi tinggi yang kompleks dari faktor-faktor yang mempengaruhi tingkat keberhasilan telemarketing. Kemudian mengambil fitur dimensi rendah yang diekstraksi sebagai inputan untuk nonlinier Support Vector Machine (SVM) sebagai pelatihan dan prediksi. Hasil menunjukkan bahwa model prediksi pemasaran memiliki kemampuan belajar yang baik dan kemampuan generalisasi yang dapat memberikan referensi pengambilan keputusan tertentu untuk bank dan industri lain untuk mencapai pemasaran presisi dengan accuracy = 86.07%, precision = 83.64%, specificity = 82.82%, recall = 89.38%. tetapi penelitian ini belum sampai pada penentuan atribut yang berpengaruh. Penelitian [8] mengusulkan pendekatan Artificial Neural Network (ANN) untuk memprediksi keberhasilan panggilan telemarketing untuk menawarkan deposito jangka panjang bank untuk memvalidasi model yang diusulkan menggunakan data pemasaran bank dari 41188 panggilan telepon. Hasil menunjukkan bahwa Artificial Neural Network (ANN) mencapai akurasi sebesar 98,93%. Tetapi jumlah data nasabah yang berlangganan deposito tidak sama dengan jumlah data nasabah yang tidak berlangganan. Perbedaan jumlah data yang sangat besar tersebut menyebabkan model yang dihasilkan akan mengarah pada nasabah yang tidak berlangganan deposito. Penelitian [9] membandingkan beberapa metode seperti decision trees, bagging, boosting, and random forests untuk melakukan klasifikasi data direct marketing campaigns. Hasil menunjukkan bahwa faktor yang sangat penting untuk hasil yang diperoleh adalah keacakan sampel bootstrap yang digunakan untuk membangun model. Kemudian, Penelitian [10] mengusulkan metode Support Mesin Vektor menggunakan algoritma Adaboost untuk mengklasifikasikan pelanggan potensial untuk pemasaran langsung. Hasil menunjukkan bahwa metode Support Mesin Vektor menggunakan algoritma Adaboost menghasilkan akurasi 95,07% dan sensitivitasnya 91,65% lebih tinggi dari pendekatan SVM biasa. Penelitian ini berhasil mengatasi permasalahan keacakan sample bootstrap Tetapi penelitian ini masih terbatas pada satu partisi data menggunakan data tidak seimbang dan menurut [11] meskipun support vector machine tunggal menunjukkan kinerja yang baik dalam klasifikasi, akan tetapi sensitif terhadap pengaturan sampel dan parameter. Oleh karena itu untuk mengatasi masalah tersebut digunakan Least Square Support Vector Machine (LSSVM) yaitu dengan menyelesaikan fungsi kendala berbentuk persamaan linear untuk mendapatkan solusi dari permasalahan quadratic programming.

III. METODE PENELITIAN

A. Support Vector Machine

Support Vector Machine digunakan untuk memecahkan masalah klasifikasi pola dan sangat cocok digunakan pada data yang dapat dipisahkan secara linear dengan memaksimalkan batas bidang pemisah (Hyperplane) untuk memisahkan data ke dalam dua class pada sebuah ruang fitur [12].

Gambar 1 menunjukkan pilihan hyperplane yang mungkin untuk set data dan bidang pemisah terbaik dengan margin maksimal yang tertelak di antara kedua kelas.



Gbr. 1 Konsep Support Vector Machine.

B. Least Square Support Vector Machine

Least Squares Support Vector Machine merupakan pengembangan SVM yang dipelopori oleh Suykens dan Vandewalle pada tahun 1999 [13]. Least Squares Support Vector Machine menghasilkan tingkat error yang lebih rendah dibandingkan dengan Support Vector Machine karena melakukan klasifikasi lanjutan yaitu dengan mengolah kembali data yang salah terklasifikasi menggunakan quadratic hyperplane. Bentuk Least Squares Support Vector Machine ditunjukkan dengan fungsi objektif berikut :

$$\min \quad \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \tag{1}$$

Dengan kendala

$$y_i(\varphi(x_i).w^T + b) = 1 - \xi_i, k = 1,2,\ldots,N \tag{2}$$

Dimana ξ adalah variabel slack yang menentukan tingkat kesalahan klasifikasi (misclassification) sampel data. C > 0 adalah parameter yang menentukan besar penalti akibat kesalahan dalam klasifikasi data. K adalah fungsi kernel yang digunakan.

C. Adaptive Boosting (AdaBoost)

Secara umum algoritma AdaBoost melatih pengklasifikasian dasar secara sekuensial dalam setiap iterasi menggunakan data latih dengan koefisien bobot yang bergantung dari performa pengklasifikasian pada iterasi sebelumnya untuk memberikan bobot yang lebih besar pada data yang salah terklasifikasi [14], [15]

langkah-langkah algoritma Adaboost adalah sebagai berikut:

1. Input: Data pelatihan beserta labelnya $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, Component Learn, dan jumlah perputaran T
2. Inisialisasi bobot data pelatihan:

$$D_i^1 = \frac{1}{N}, i = 1, \ldots, N \tag{3}$$

3. Untuk iterasi *t=1,...,T*
   a. Gunakan algoritma Component Learner untuk melatih suatu komponen klasifikasi, h_t pada bobot pelatihan.
   b. Hitung bobot kesalahan klasifikasi pada h_t
   c.

$$\epsilon_t = \sum_{i=1}^{N}(D_i^t(y_i \neq h_t(x))) \tag{4}$$

   Indeks kepercayaan pembelajaran dihitung sebagai:

$$c_t = \frac{1}{2}ln\left(\frac{(1-\epsilon_t)}{\epsilon_t}\right) \tag{5}$$

   d. Perbaharui bobot sampel pelatihan

$$D_i^{t+1} = \frac{D_i^t exp\ (c_t x y_i \neq y_i^*)}{\sum_{i=1}^{N} D_i^t} \tag{6}$$

   e. Testing model dengan data uji
4. Keluaran pembelajaran terakhir
   Kombinasi semua klasifikasi

$$h_i = sign\left(\sum_{t=1}^{T} c_t h_t(x)\right) \tag{7}$$

D. Ensemble Least Square Support Vector Machine dengan Algoritma Adaboost

a) Load Dataset
b) Identifikasi Label Atribut, class dan jumlah data.
c) Menentukan jumlah data latih dan data uji.
d) Membentuk klasifikasi
- *Menentukan Nilai Gamma.*
- *Menentukan Nilai Sigma.*
- *Inisialisasi Variabel, dan bobot awal.*

e) Menentukan jumlah perulangan (Iterasi)
- *Menggunakan algoritma component learner (least squares support vector machine) untuk membentuk suatu model klasifikasi.*
- *Menghitung bobot kesalahan klasifikasi*
- *Memperbaharui bobot sampel data latih*
- *Pengujian model dengan data uji*
- *Hasil dari Pengujian model.*

E. Evaluasi Kinerja

Evaluasi hasil Kinerja dari setiap klasifikasi di hitung berdasarkan 3 pengukuran yaitu : Accuracy, Sensitivity, dan Specificity berdasarkan nilai True Positive, False Negative, False Positive, dan True Negative [16].

1. True Positive (TP) : Jumlah data yang teridentifikasi sebagai 'Pelanggan Potensial' dan hasil prediksinya 'Pelanggan Potensial'.
2. True Negative (TN) : Jumlah data yang teridentifikasi sebagai 'Pelanggan Non-Potensial' dan hasil prediksinya 'Pelanggan Non-Potensial'.
3. False Positive (FN) : Jumlah data yang teridentifikasi sebagai 'Pelanggan Non-Potensial' tetapi hasil prediksinya 'Pelanggan Potensial'.
4. False Negative (FN) : Jumlah data data yang teridentifikasi sebagai 'Pelanggan Potensial' tetapi hasil prediksinya 'Pelanggan Non-Potensial'.

$$Accuracy = \frac{TP+TN}{TN+FP+FN+TP} \tag{8}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{9}$$

$$Specificity = \frac{TN}{TN+FP} \quad (10)$$

TABEL I
CONFUSION MATRIX.

Prediksi	Pelanggan Potensial	Pelanggan Non-Potensial
Pelanggan Potensial	True Positive (TP)	False Negative (FN)
Pelanggan Non-Potensial	False Positive (FP)	True Negative (TN)

IV. HASIL DAN PEMBAHASAN

Data yang diolah adalah data bank direct marketing dataset yang dapat diakses melalui University of California at Irvine (UCI) Machine Learning Repository. Data tersebut berasal dari bank Portugal, dari bulan mei 2008 sampai bulan juni 2013 dengan jumlah total nasabah 41.118. pemilihan dataset tersebut karena relevansi dengan konteks, jumlah nasabah yang besar, periode data yang luas yakni 2008 s.d 2013, dan keberagaman variabel. Deskripsi dataset ditunjukkan pada Tabel 2.

TABEL III
DESKRIPSI DATASET.

Atribut	**Tipe data**
Age	Numeric
Job	Categorical Admin, Blue-collar, Enterpreneur, Housemaid, Management, Retired, Self-employed, Services, Student, Technician, Unemployed, Unknown
Martial	Categorical Divorced, Married, Single, Unknown
Education	Categorical : Basic.4y, Basic.6y, Basic.9y, High.school, Illiterate, Professional.course, University.degree, Unknown
Default	Categorical : No, Yes, Unknown
Housing	Categorical : No, Yes, Unknown
Loan	Categorical : No, Yes, Unknown
Contact	Categorical : Celular, Telephone
Month	Categorical Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec
Day_of_week	Categorical Mon, Tue, Wed, Thu, Fri
Duration	Numeric
Campaign	Numeric
Pdays	Numeric
Previous	Numeric
Poutcome	Categorical Failure, Nonexistent, Success
Y	Binary : Yes, No

Dari jumlah total data keseluruhan nasabah bank Portugal, dari bulan mei 2008 sampai bulan juni 2013 sebanyak 41.188 nasabah, dinormalisasi menjadi 9.280 untuk menyeimbangkan antara nasabah pelanggan deposito berjangka dan bukan nasabah pelanggan deposito berjangka. Hal ini disebabkan jumlah data nasabah yang berlangganan deposito tidak sama dengan jumlah data nasabah yang tidak berlangganan. perbandingan jumlah data yang nasabah yang berlangganan deposito sebanyak 4,640 data, sedangkan nasabah yang tidak berlangganan sebanyak 36,540 data. Dengan perbedaan jumlah data yang sangat besar tersebut menyebabkan model yang dihasilkan akan cenderung mengarah pada nasabah yang tidak berlangganan deposito. Sehingga model tersebut tidak efisien untuk digunakan. Jadi data yang diolah sebanyak 4,640 data pelanggan deposito berjangka dan 4,640 data bukan pelanggan deposito berjangka.

Penelitian ini mengusulkan untuk melihat kinerja metode Ensemble Least Square Support Vector Machine dengan pembagian data latih sebesar 7424 dan data uji sebesar 1856. Pemilih data yang menjadi data latih berdasarkan urutan waktu data dikumpulkan. Misalnya dari 9280 total data yang diolah, maka 7424 data awal yang didapatkan yang akan menjadi data latih dan 1856 data sisanya yang akan menjadi data uji. Sebelum melakukan pelatihan terlebih dahulu diperlukan pemilihan atribut dengan pemilihan fitur information gain karena pada pelatihan menggunakan support vector machine tidak terdapat mekanisme untuk mengetahui atribut penting atau yang kurang penting. Atribut yang kurang penting umumnya tidak mempengaruhi efektifitas teknik klasifikasi. Oleh karena itu, jika atribut yang kurang penting ini dibuang maka efisiensi teknik klasifikasi akan meningkat. Fokus dari penelitian ini untuk menerapkan metode ensemble least squares support vector machine menggunakan AdaBoost.

Sebelum proses klasifikasi terlebih dahulu kami akan menampilkan kinerja dari model yang dihasilkan. Dalam analisis kinerja berbagai model, ditemukan bahwa model Support Vector Machine mencapai akurasi tertinggi sebesar 98,91%. Diikuti oleh model Least Square Support Vector Machine dengan akurasi 98,16%, dan model Ensemble Least Square Support Vector Machine yang menggunakan teknik Adaboost juga mencapai akurasi yang sama, yaitu 98,16%. Adapun kinerja dari model yang dihasilkan ditunjukkan pada Tabel 3.

TABEL IIIII
HASIL KINERJA MODEL.

Metode	Akurasi Model (%)
Support Vector Macine	98,91
Least Square Support Vector Machine	98,16
Ensemble Least Square Support Vector Machine Menggunakan Adaboost	98,16

Proses pengujian menggunakan software Matlab untuk melihat kinerja dari metode – metode yang diterapkan. Pengujian metode yang diterapkan menggunakan dua pengujian data yaitu ketika jumlah data tidak seimbang antara data latih dan data uji serta jumlah data seimbang. Berdasarkan hasil confusion matrix seperti yang terlihat pada Tabel 4. Sedangkan hasil kinerja dari masing-masing metode ditampilkan pada Tabel 5.

TABEL IVV
HASIL CONFUSION MATRIX

Metode	TP	TN	FP	FN
Support Vector Macine	908	841	20	87
Least Square Support Vector Machine	906	852	22	76
Ensemble Least Square Support Vector Machine Menggunakan Adaboost	907	859	21	69

TABEL V
HASIL KINERJA METODE.

Metode	Accuracy %	Sensitivity %	Specificity %
Support Vector Macine	94,23	91,26	97,68
Least Square Support Vector Machine	94,72	92,26	97,48
Ensemble Least Square Support Vector Machine Menggunakan Adaboost	95,15	92,93	97,61

Hasil yang didapatkan metode ensemble least square support vector machine menggunakan adaboost mendapatkan persentase yang paling tinggi diantara metode support vector machine dan least square support vector machine. Hal ini membuktikan bahwa metode ensemble least square support vector machine menggunakan adaboost lebih baik dibandingkan dengan metode lainnya. Metode Least Squares Support Vector Machine menunjukkan kinerja yang lebih baik dibandingkan dengan metode Support Vector Machine berdasarkan hasil klasifikasi yang didapatkan. Metode yang diusulkan yakni Ensemble Least Squares Support Vector Machine berhasil meningkatkan kinerja dari Least Squares Support Vector Machine karena proses pembobotan ulang terhadap data yang salah terklasifikasi dan melakukan pelatihan ulang pada setiap iterasi hingga batas iterasi yang ditentukan. Akan tetapi, semakin banyak iterasi yang dilakukan oleh metode ini, tidak akan menjamin kinerja dari component learner akan semakin baik dan ini akan menjadi pengembangan selanjutnya.

V. KESIMPULAN

Metode Ensemble Least Squares Support Vector Machine yang menggunakan teknik Adaboost berhasil meningkatkan kinerja dalam mengklasifikasikan pelanggan deposito potensial dalam dataset dari bank Portugal. Hasil penelitian menunjukkan bahwa metode ini mencapai akurasi sebesar 95.15%, sensitivitas sebesar 92.93%, dan spesifisitas sebesar 97.61%. Dalam perbandingan dengan model Support Vector Machine dan Least Square Support Vector Machine tunggal, Ensemble Least Squares Support Vector Machine mengungguli keduanya dalam hal akurasi. Selain itu, penelitian ini juga mengungkapkan bahwa pemilihan fitur menggunakan metode Information Gain dan pembagian data latih serta data uji berdasarkan urutan waktu pengumpulan data memainkan peran penting dalam meningkatkan kinerja model. Proses ini membantu mengurangi jumlah atribut yang kurang penting dan memastikan bahwa data latih dan data uji memiliki karakteristik yang representatif. Namun, penelitian ini juga mengidentifikasi bahwa jumlah iterasi dalam metode Ensemble Least Squares Support Vector Machine perlu dipertimbangkan dengan cermat, karena peningkatan iterasi tidak selalu menghasilkan kinerja yang lebih baik. Oleh karena itu, pengembangan selanjutnya dapat difokuskan pada penelitian tentang pengaruh jumlah iterasi pada metode ini. Secara keseluruhan, penelitian ini memberikan kontribusi dalam pengembangan metode klasifikasi yang efisien dalam mengidentifikasi pelanggan deposito potensial, yang memiliki implikasi potensial dalam strategi pemasaran dan pengambilan keputusan di industri perbankan.

UCAPAN TERIMA KASIH

Terima kasih disampaikan kepada Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, Direktoral Jenderal Perguruan Tinggi yang telah memberikan pendanaan melalui hibah penelitian tahun 2023 dan pihak-pihak yang telah mendukung terlaksananya penelitian ini.

DAFTAR PUSTAKA

[1] Q. Zhuang and Y. Yao, “Application of data mining in term deposit marketing,” iaeng.org, 2018, Accessed: Mar. 08, 2022. [Online]. Available: http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp707-710.pdf.

[2] T. Parlar and A. SK, “Using data mining techniques for detecting the important features of the bank direct marketing data,” Int. J. Econ. Financ. Issues, vol. 7, no. 2, pp. 692–696, 2017, Accessed: Mar. 08, 2022. [Online]. Available: https://dergipark.org.tr/en/pub/ijefi/issue/32035/354551?publisher=http-www-cag-edu-tr-ilhan-ozturk.

[3] P. Ruangthong, S. J.-2015 12th I. Joint, and undefined 2015, “Bank direct marketing analysis of asymmetric information based on machine learning,” ieeexplore.ieee.org, Accessed: Mar. 08, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7219777/.

[4] J. Asare-Frempong and M. Jayabalan, “Predicting customer response to bank direct telemarketing campaign,” ieeexplore.ieee.org, 2017, doi: 10.1109/ICE2T.2017.8215961.

[5] O. Apampa, “Evaluation of classification and ensemble algorithms for bank customer marketing response prediction,” J. Int. Technol. Inf. Manag., vol. 25, 2016,

Accessed: Mar. 08, 2022. [Online]. Available: https://scholarworks.lib.csusb.edu/jitim/vol25/iss4/6/.

[6] A. Mutoi Siregar, S. Faisal, H. H. Handayani, and A. Jalaludin, "Classification Data for Direct Marketing using Deep Learning," Sci. Jounal PPI-UKM, vol. 7, no. 2, 2020, doi: 10.27512/sjppi-ukm/se/a15052020.

[7] J. Che, S. Zhao, and Y. Li, "Bank telemarketing forecasting model based on t-SNE-SVM," scirp.org, 2020, Accessed: Mar. 08, 2022. [Online]. Available: https://www.scirp.org/journal/paperinformation.aspx?paperid=100260.

[8] M. Selma, "Predicting the success of bank telemarketing using Artificial Neural Network," Int. J. Econ., 2020, Accessed: Mar. 08, 2022. [Online]. Available: https://publications.waset.org/10010974/predicting-the-success-of-bank-telemarketing-using-artificial-neural-network.

[9] D. Grzonka, G. Suchacka, B. B.-I. S. in, and undefined 2016, "Application of selected supervised classification methods to bank marketing campaign," yadda.icm.edu.pl, vol. 5, no. 1, pp. 36–48, 2016, Accessed: Mar. 08, 2022. [Online]. Available: https://yadda.icm.edu.pl/yadda/element/bwmeta1.element.desklight-3c731462-b2de-4c6d-9e43-95ccf418785f.

[10] A. Lawi, A. A. Velayaty, and Z. Zainuddin, "On Identifying Potential Direct Marketing Consumers using Adaptive Boosted Support Vector Machine," in Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology, CAIPT 2017, Aug. 2018, pp. 1–4, doi: 10.1109/CAIPT.2017.8320691.

[11] L. Zhou and K. Lai, "Least squares support vector machines ensemble models for credit scoring," Elsevier, 2010, Accessed: Mar. 08, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417409004394.

[12] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," Data Min. Knowl. Discov., vol. 2, no. 2, pp. 121–167, 1998, doi: 10.1023/A:1009715923555.

[13] J. Suykens and J. V. Letters, "Least squares support vector machine classifiers," Springer, vol. 9, pp. 293–300, 1999, Accessed: Mar. 08, 2022. [Online]. Available: https://link.springer.com/article/10.1023/A:1018628609742.

[14] F. Schwenker, "Ensemble methods: Foundations and algorithms," ieeexplore.ieee.org, 2022, Accessed: Mar. 08, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6410720/.

[15] R. Schapire, "Boosting: Foundations and algorithms," emerald.com, 2013, Accessed: Mar. 08, 2022. [Online]. Available: https://www.emerald.com/insight/content/doi/10.1108/03684921311295547/full/html.

[16] R. Kohavi, "Guest editors' introduction: On applied research in machine learning."