

# Peningkatan performance Logistic Regression menggunakan teknik Ensemble Bagging pada kasus Credit Scoring

Firman Aziz  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Pancasakti  
Makassar, Indonesia  
firman.aziz@unpacti.ac.id

## Abstrak

Abstrak Model credit scoring menjadi sangat efektif untuk dikembangkan karena jumlah data pelanggan yang sangat besar dalam sebuah industry kredit. Kinerja prediktif terbaik untuk penyelesaian masalah credit scoring menggunakan model statistic. Tetapi keakuratannya dapat ditingkatkan menggunakan persamaan nonlinear dengan memperkirakan parameter dari suatu model. Penelitian ini mengusulkan penerapan metode Ensemble Logistic Regression untuk menyelesaikan masalah credit scoring. Parameter Newton Raphson dari metode Logistic Regression di ensemble menggunakan teknik bagging. German dan Australian Dataset akan diolah untuk melihat kinerja dari metode yang diusulkan. Hasil menunjukkan bahwa metode yang diusulkan berhasil diterapkan dalam credit scoring dengan akurasi sebesar 79.6 % untuk German Dataset dan 86.9 % untuk Australian Dataset.

Kata Kunci: *Credit Scoring, Logistic Regression, Ensemble, Bagging, Newton Raphson.*

## Abstract

*Abstract The credit scoring model is very effective to develop because of the huge amount of customer data in a credit industry. The best predictive performance for solving credit scoring problems using a statistical model. But its accuracy can be improved using nonlinear equations by estimating the parameters of a model. This study proposes the application of the Ensemble Logistic Regression method to solve credit scoring problems. Newton Raphson parameters of the Logistic Regression method are ensembled using bagging techniques. German and Australian Dataset will be processed to see the performance of the proposed method. The results show that the proposed method was successfully applied in credit scoring with an accuracy of 79.6% for German Dataset and 86.9% for Australian Dataset.*

*Keywords: Credit Scoring, Logistic Regression, Ensemble, Bagging, Newton Raphson.*

## 1. Pendahuluan

Kredit adalah kegiatan bisnis yang banyak dilakukan oleh bank dan menjadi sumber penghasilan terbesar sekaligus risiko terbesar. Oleh karena itu stabilitas bank sangat dipengaruhi oleh keberhasilan dalam mengelola kredit. Credit Scoring adalah teknologi yang berupaya meminimalkan risiko peminjam kredit yang tidak dapat melakukan pembayaran tepat waktu. Credit Scoring adalah masalah klasifikasi yang

membagi dua calon peminjam menjadi peminjam yang disetujui dan peminjam yang tidak disetujui berdasarkan beberapa karakteristik seperti usia, kondisi ekonomi, status sosial, jaminan, dll (Kim & Sohn, 2004).

Dalam proses klasifikasi, metode statistik banyak digunakan dalam industri penilaian kredit karena relatif lebih mudah dipahami dan diimplementasikan. Metode statistik yang paling banyak digunakan adalah Logistic Regression. Penelitian (Desai et al., 1996) memperkirakan jaringan saraf tiruan dan mengusulkan model dalam pemeringkatan kredit. Model Regresi Logistik memiliki persentase kinerja yang lebih baik daripada metode lain dalam mengklasifikasikan peminjam baik dan buruk. Penelitian (Lessmann et al., 2015) mempelajari tentang algoritma klasifikasi terbaru pada aplikasi kasus kredit. Hasil menunjukkan bahwa pada kasus Credit Scoring, metode Logistic Regression memberikan kinerja yang lebih baik.

Ensemble merupakan beberapa set pelatihan yang digunakan untuk memecahkan masalah yang sama dari hasil dari klasifikasi tunggal kemudian digabungkan dengan teknik Ensemble menjadi pengelompokan tunggal untuk meningkatkan kinerja dari klasifikasi tunggal (Tsai, 2014). Penelitian (Devi & Chezian, 2016) mengusulkan untuk menganalisis keakuratan metode Ensemble dalam mengklasifikasikan pelanggan menggunakan tiga metode Ensemble yaitu AdaBoost, Bagging, dan Random Forest tetapi dalam penelitian ini masih menggunakan klasifikasi standar tunggal dari metode Ensemble yaitu metode Decision Tree.

Penelitian (Alam et al., 2015) menerapkan metode pemilihan fitur pada German dataset dan memasukkan klasifikasi tunggal dengan metode greedy stepwise search method tetapi penelitian ini mengurangi atribut dari 20 menjadi 14. Penelitian (West et al., 2005) melakukan Ensemble bagging dan boosting dengan menggunakan JST sebagai klasifikasi tunggal. Ensemble ANN menunjukkan hasil yang akurat dan kesalahan generalisasi yang rendah tetapi Ensemble bagging memiliki kelemahan ketika jumlah atribut dan data sedikit. Penelitian (Nanni & Lumini, 2009) mengkaji penggunaan SVM dan KNN sebagai klasifikasi tunggal dan menggunakan Ensemble bagging dan boosting untuk meningkatkan akurasi klasifikasi tunggal. Hasilnya menunjukkan keunggulan model Ensemble dalam hal akurasi. Penelitian (Ala'raj & Abbod, 2015) mengusulkan model Ensemble homogen dan heterogen sistematis berdasarkan tiga klasifikasi LR, JST dan SVM. Hasil penelitian menunjukkan bahwa pengelompokan Ensemble heterogen memberikan akurasi prediksi tinggi dengan kesalahan rendah daripada pengelompokan homogen dan tunggal tetapi dalam penelitian ini, tidak ada estimasi yang diperoleh pada Ensemble dengan teknik bagging. Penelitian (Lawi et al., 2017) mengusulkan Ensemble Logistic Regression dengan metode Boosting menggunakan Algoritma GradientBoost pada kasus Credit Scoring. Dua set data diimplementasikan yaitu German dan Australian dataset. Hasil penelitian menunjukkan bahwa Ensemble GradientBoost berhasil meningkatkan kinerja klasifikasi Logistic Regression dan mencapai tingkat akurasi tertinggi di kedua dataset. Penelitian (Dahiya et al., 2015) membagi tiga partisi data untuk melihat kinerja chi-square. Fokus dari penelitian ini terletak pada pemilihan fitur dan penelitian menunjukkan bahwa akurasi kinerja klasifikasi diperoleh pada partisi data dengan 70% untuk data pelatihan dan 30% data uji.

Dalam penelitian ini, diusulkan metode Ensemble Logistic Regression dengan teknik bagging untuk kasus Credit Scoring. Teknik bagging digunakan karena mengurangi masalah overfitting dan tidak tergantung pada teknik pengklasifikasi tunggal dan bagging lebih efektif pada algoritma pembelajaran yang tidak stabil. Fokus dari penelitian ini adalah memperkirakan Ensemble Logistic Regression menggunakan

parameter Newton Raphson. Newton Raphson dipilih karena dapat memecahkan tipe data nonlinier dengan menafsirkan nilai awal untuk fungsi maksimumnya.

## 2. Metodologi

### 2.1. Logistic Regression

Klasifikasi tunggal dalam penelitian ini menggunakan Logistic Regression yang merupakan metode statistik paling banyak digunakan untuk menyelesaikan masalah klasifikasi dan regresi (Akkoc, 2012). Logistic Regression digunakan untuk membuat contoh variabel hasil biner, biasanya diwakili oleh 1 atau 0. Skor model menggunakan angka biner di mana 1 untuk debitur baik dan 2 untuk debitur buruk dan setiap variabel independen memberikan ketergantungan pada setiap skor (Thomas, 2000). Fungsi classifier ditunjukkan oleh persamaan:

$$\log\left(\frac{p}{1-p}\right) = \sum_{i=1}^n \beta^{(i)} * x^{(i)} + e = \beta^T x + e \quad (1)$$

Di mana  $\beta = (\beta^{(1)}, \beta^{(2)}, \dots, \beta^{(n)})$  adalah vektor koefisien hyperplane. Probabilitas pelanggan berhenti pada persamaan (1) dapat dengan mudah dirumuskan seperti pada persamaan berikut (2).

$$p = \frac{e^{\beta^T x + e}}{1 + e^{\beta^T x + e}} \quad (2)$$

Untuk memperoleh estimasi parameter Logistic Regression dapat dilakukan dengan dua cara yaitu dengan Maximum Likelihood Estimation (MLE) dan Newton Raphson. Namun, penelitian ini menggunakan parameter Newton Raphson.

Newton Rhapsion adalah metode untuk menyelesaikan persamaan nonlinear seperti memecahkan persamaan likelihood dalam model Logistic Regression (Agresti, 2010). Metode Newton Rhapsion membutuhkan estimasi awal dari nilai fungsi maksimumnya, yang merupakan estimasi menggunakan pendekatan polinomial dua derajat.

Dalam hal ini untuk menentukan nilai  $\hat{\beta}$  dari  $\beta$  yang merupakan fungsi maksimum  $g(\beta)$  Misalkan  $q' = \left(\frac{\partial g}{\partial \beta_1}, \frac{\partial g}{\partial \beta_2}, \dots\right)$ , dan anggap  $\mathbf{H}$  dilambangkan sebagai matriks dengan anggota  $h_{ab} = \frac{\partial^2 g}{\partial \beta_1 \partial \beta_2}$  anggaplah  $q^{(t)}$  dan  $\mathbf{H}^{(t)}$  adalah bentuk evaluasi estimasi  $\beta^{(t)}$  ke  $t$  pada  $\hat{\beta}$ . Pada langkah  $t$  dalam proses iterasi ( $t = 0, 1, 2, \dots$ ).

### 2.2. Metode Ensemble

Fokus dari klasifikasi ini adalah untuk memecahkan masalah serupa dengan menggabungkan satu set klasifikasi untuk mendapatkan klasifikasi yang lebih akurat (Zhou et al., 2010). Metode ensemble dapat mengurangi kesalahan klasifikasi secara efektif, dan diyakini memiliki kinerja baik dibandingkan dengan penggunaan klasifikasi tunggal.

Dibandingkan dengan klasifikasi tunggal yang belajar dan melatih satu data saja, klasifikasi ensemble melakukan pembelajaran dan pelatihan berbagai data dan hasil

pelatihan akan membangun hipotesis dan menghasilkan akurasi yang lebih baik (Kuncheva, 2014). Klasifikasi tunggal dalam penelitian ini menggunakan Logistic Regression yang merupakan metode statistik paling banyak digunakan untuk menyelesaikan masalah

Beberapa teknik pengklasifikasi ansambel telah dikembangkan seperti random forest, rotation forest, boosting, and bagging. karena fokus dari penelitian ini adalah menggunakan ensemble bagging, maka klasifikasi ensemble bagging akan disajikan.

Bagging adalah algoritma yang memberikan kinerja yang baik dan sangat mudah diterapkan. Jumlah *bag* yang berbeda (*n-bag*) yang diperoleh melalui data pelatihan dan hasil dataset pelatihan pada setiap *bag* akan diisi dengan data yang dihasilkan secara acak dari dataset pelatihan (*Ala'raj M dan Abbod M*). Bagging (Bootstrap Aggregating) membuat sampel *M* bootstrap  $T_1, T_2, \dots, T_M$  diambil secara acak dari set pelatihan asli *T* ukuran *n* (*Ala'raj M dan Abbod M*). Setiap sampel bootstrap  $T_i$  ukuran *n* kemudian digunakan untuk melatih classifier dasar  $C_i$ . Prediksi pada pengamatan baru dibuat dengan mengambil suara mayoritas dari ensemble  $C^*$ . dibangun dari  $C_1, C_2, \dots, C_M$  (Marques et al., 2012). Algoritma ensemble bagging sebagai berikut:

satu algoritma klasifikasi tunggal  $C_i(x)$  akan diberikan satu set pelatihan ukuran *n*.

- a. Masukan urutan sampel pelatihan  $(x_1:y_1), \dots, (x_n:y_n)$  dengan label  $y \in Y = \{-1, 1\}$
- b. Inisialisasi probabilitas untuk setiap contoh dalam set pembelajaran  $D_1 (i) = 1/n$  and set  $t = 1$ .
- c. Loop dimana  $t < B = 100$  anggota ensemble
  - Bentuk pelatihan set ukuran *n* dengan pengambilan sampel penggantian dari distribusi  $D_t$ .
  - Dapatkan hipotesis  $h_t: X \rightarrow Y$
  - Set  $t = t + 1$

Akhir dari loop

- d. Keluarkan hipotesis akhir ensemble.

$$C^*(x_i) = h_{final}(x_1) = \operatorname{argmax} \sum_{t=1}^B I(C_t(x) = y).$$

### 2.3. Evaluasi Kinerja

Evaluasi kinerja metode klasifikasi dapat dilihat dari tingkat kesalahan klasifikasi. Untuk menghitung tanda kesalahan klasifikasi dapat menggunakan *confusion matrix*, biasanya disebut dengan tabel kontingensi seperti pada Tabel 1.

Tabel 1. Confusion Matrix.

Prediksi	Peminjam Baik	Peminjam Buruk
Peminjam Baik	True Positive (TP)	False Negative (FN)
Peminjam Buruk	False Positive (FP)	True Negative (TN)

Tingkat akurasi didefinisikan sebagai tingkat kedekatan antara data yang diprediksi dengan data aktual atau rasio jumlah data yang diklasifikasikan dengan benar sebagaimana ditunjukkan oleh persamaan (3)

$$Akurasi = \frac{TP+TN}{TN+FP+FN+TP} \quad (3)$$

### 3. Eksperimental

#### 3.1. Dataset

Untuk mengevaluasi keakuratan model yang diusulkan digunakan dua data yaitu German dan Australian dataset. dalam literatur penelitian kredit, kedua dataset ini digunakan untuk menguji model dan tersedia dalam UCI machine learning repository. Secara umum, dataset dapat dilihat pada Tabel 2.

Tabel 2. Deskripsi Dataset.

Dataset	Jumlah Peminjam	Peminjam Baik	Peminjam Buruk	Jumlah Atribut
German dataset	1000	700	300	20
Australian dataset	690	307	383	14

Kedua dataset akan dibagi menjadi 70% pelatihan data dan 30% pengujian data untuk melihat hasil klasifikasi model yang diusulkan.

#### 3.2. Normalisasi Data

Data harus diubah dari skala nilai yang berbeda untuk menghilangkan redundansi dan membangun model data pada interval yang sama. Atribut dataset dinormalisasi dengan nilai 0 untuk peminjam 'buruk' dan 1 untuk peminjam 'baik'.

#### 3.3. Implementasi

Fokus dari penelitian ini adalah untuk melihat penggunaan parameter Newton Raphson dari klasifikasi Ensemble Logistic Regression karena penggunaan parameter Newton Raphson dapat meningkatkan kinerja Ensemble Logistic Regression. Logistic Regression akan digunakan sebagai klasifikasi tunggal dan teknik ensemble yang digunakan adalah bagging.

### 4. Hasil

Untuk melihat kinerja metode yang diusulkan menggunakan proses pengujian dengan bahasa pemrograman python. Tingkat akurasi didasarkan pada parameter Newton Raphson dari Ensemble Logistic Regression.

Tabel 3. Hasil Klasifikasi.

Klasifikasi	German dataset	Australian dataset
Logistic Regression	77 %	85.9 %

Ensemble Logistic Regression	79.6 %	86.9 %
---------------------------------	--------	--------

---

Tabel 3 menunjukkan hasil klasifikasi pada german dan australian dataset. Metode Logistic Regression menghasilkan akurasi sebesar 77,0% pada German dataset dan Australian dataset menghasilkan akurasi sebesar 85,9%. Hasil dari model yang diusulkan yaitu Ensemble Logistic Regression menggunakan parameter Newton Raphson berhasil mencapai kinerja terbaik dengan meningkatkan kinerja klasifikasi tunggal dengan akurasi 79,6% untuk german dataset dan 86,9% dari dataset Australia.

dari hasil yang diperoleh, peningkatannya tidak terlalu signifikan karena australian dataset hanya meningkat sebesar 1% dan german dataset hanya meningkat sebesar 2,6%. tetapi dalam penelitian ini diperoleh pengetahuan baru yaitu ketika jumlah iterasi 10 hingga 50 akan mempengaruhi keakuratan akurasi klasifikasi Ensemble Logistic Regression tetapi jika iterasi > 50 nilai akurasi akan menjadi konstan. Selain itu, didapatkan pula bahwa metode yang diusulkan sensitif terhadap jumlah data yang kecil.

## 5. Kesimpulan

Dalam penelitian ini diusulkan metode Ensemble Logistic Regression dengan teknik Bagging untuk kasus Credit Scoring. Fokus dari penelitian ini adalah meningkatkan kinerja Logistic Regression menggunakan parameter Newton Raphson kemudian dilakukan teknik ensemble menggunakan bagging.

Metode yang diusulkan berhasil mencapai kinerja terbaik dari klasifikasi tunggal. untuk akurasi german dataset menghasilkan 79,6% dan 86,9% australian dataset. Jumlah iterasi 10 hingga 50 mempengaruhi akurasi klasifikasi tetapi ketika iterasi > 50 nilai akurasi konstan dan dalam penelitian ini metode yang diusulkan sensitif terhadap jumlah data kecil.

## Daftar Pustaka

- Kim, Y. S., & Sohn, S. Y. (2004). Managing loan customers using misclassification patterns of credit scoring model. *Expert Systems with Applications*, 26(4), 567-573.
- Desai, V. S., Crook, J. N., & Overstreet Jr, G. A. (1996). A comparison of neural networks and linear scoring models in the credit union environment. *European journal of operational research*, 95(1), 24-37.
- Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, 247(1), 124-136.
- Tsai, C. F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16, 46-58.
- Devi, C. D., & Chezian, R. M. (2016, October). A relative evaluation of the performance of ensemble learning in credit scoring. In *2016 IEEE International Conference on Advances in Computer Applications (ICACA)* (pp. 161-165). IEEE.
- Alam, L., & Mamun, T. I. (2015). An Analytical Comparison on Filter Feature Extraction Method in Data Mining using J48 Classifier. *International Journal of Computer Applications*, 975, 8887.
- West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. *Computers & operations research*, 32(10), 2543-2559.

- Nanni, L., & Lumini, A. (2009). An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. *Expert systems with applications*, 36(2), 3028-3033.
- Ala'raj, M., & Abbod, M. (2015, September). A systematic credit scoring model based on heterogeneous classifier ensembles. In *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-7). IEEE.
- Lawi, A., Aziz, F., & Syarif, S. (2017, August). Ensemble GradientBoost for increasing classification accuracy of credit scoring. In *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)* (pp. 1-4). IEEE.
- Dahiya, S., Handa, S. S., & Singh, N. P. (2015). Credit scoring using ensemble of various classifiers on reduced feature set. *Industrija*, 43(4).
- Akkoc, S. (2012). An empirical comparison of conventional techniques, neural networks and the three stage hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. *European Journal of Operational Research*, 222(1), 168-178.
- Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International journal of forecasting*, 16(2), 149-172.
- Agresti, A. (2010). *Analysis of ordinal categorical data* (Vol. 656). John Wiley & Sons.
- Zhou, L., Lai, K. K., & Yu, L. (2010). Least squares support vector machines ensemble models for credit scoring. *Expert Systems with Applications*, 37(1), 127-133.
- Kuncheva, L. I. (2014). *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons.
- Marques, A. I., García, V., & Sánchez, J. S. (2012). Exploring the behaviour of base classifiers in credit scoring ensembles. *Expert Systems with Applications*, 39(11), 10244-10250.