



## Model Algoritma Boosted Gradient Trees untuk Penentuan Kelayakan Pemberian Kredit Koperasi

Aprinia Handayani

Universitas Pamulang

dosen@unpam.ac.id<sup>1</sup>

### Kata kunci:

Analisa kredit, pohon keputusan, algoritma Gradient Boosted Trees

### Abstrak

Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Sebelum koperasi menyetujui kredit yang diajukan oleh debitur, koperasi melakukan analisis kredit terhadap debitur apakah pengajuan kredit disetujui atau tidak disetujui. Analisa kredit yang dilakukan oleh analis terkadang tidak akurat, sehingga beberapa debitur yang diberikan kredit tidak mempunyai kemampuan untuk membayar yang menyebabkan kredit macet. Dari permasalahan ini diperlukan suatu model yang mampu mengklasifikasikan sekaligus memprediksi debitur mana saja yang bermasalah dan tidak bermasalah. Model pohon keputusan ternyata mampu meningkatkan akurasi dalam menganalisa kelayakan kredit yang diajukan calon debitur. Semakin kaya informasi atau pengetahuan yang dikandung oleh data training, maka akurasi pohon keputusan akan semakin meningkat.

### Pendahuluan

Koperasi adalah merupakan usaha kekeluargaan dengan tujuan untuk mensejahterakan anggotanya (UUD 1945 pasal 33 ayat 1). Umumnya koperasi dikendalikan secara bersama oleh seluruh anggotanya dimana setiap anggota memiliki hak suara yang sama dalam setiap keputusan yang diambil. Ada beberapa tahap yang harus dilalui sebelum pinjaman dicairkan, mulai dari calon debitur mengisi data dan mengajukan aplikasi, setelah aplikasi diterima oleh koperasi, maka koperasi akan melakukan survey untuk memastikan apakah data yang diisi oleh calon debitur benar adanya. Apabila tidak ada masalah dalam proses survey, koperasi akan melakukan analisa berdasarkan aplikasi calon debitur, termasuk latar belakang, kemampuan debitur dan lain-lain. Jika tahap ini sudah disetujui maka dana yang diajukan oleh debitur akan dicairkan).

Menurut Pasal 1 angka 11 Undang-Undang Nomor 10 Tahun 1998, kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank atau instansi keuangan dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan jumlah bunga.

Kredit yang diajukan oleh debitur memiliki resiko, karena dari sekian banyak debitur yang mengajukan aplikasi ada kemungkinan beberapa debitur yang bermasalah dalam pembayarannya sehingga menyebabkan kredit macet. Sebelum koperasi menyetujui kredit

yang diajukan oleh debitur, koperasi melakukan analisis kredit terhadap debitur apakah pengajuan kredit disetujui atau tidak disetujui. Komponen yang mempengaruhi risiko kredit, adalah kemungkinan debitur akan gagal membayar dalam memenuhi kontrak pembayaran, klaim yang akan ditanggung debitur jika tidak memenuhi kewajiban membayar dan nominal yang hilang akibat risiko default atau gagal bayar (Menarianti, 2015).

Analisis kredit merupakan hal yang penting dalam lingkup resiko keuangan (Lai, Lean Yu, & Zhou, 2006) oleh karena itu perlunya dilakukan analisa. Namun, melakukan proses analisa kredit membutuhkan waktu lama (Kotsiantis, Kanellopoulos, Karioti, & Tampakas, 2009) dan mengidentifikasi data debitur yang bermasalah merupakan hal yang sulit (Odeh, Featherstone, & Das, 2010). Kesalahan analisa kredit dapat menyebabkan risiko kredit, seperti hilangnya nasabah, ketidakpastian pembayaran dan pinjaman bahkan ketidakmampuan nasabah dalam mengembalikan pinjaman dana kredit. Untuk melindungi dana kredit, digunakan jaminan yang harus disediakan oleh pihak nasabah sebagai beban nasabah. Pemberian kredit dengan jaminan dapat berupa: jaminan benda berwujud (tanah, bangunan, kendaraan bermotor, kebun, perhiasan dan lain-lain), jaminan tidak berwujud (sertifikat tanah, sertifikat saham, sertifikat obligasi, SK pengangkatan kerja dan lain-lain) dan jaminan orang (jaminan yang diberikan oleh seseorang yang menyatakan kesanggupan untuk menanggung segala resiko apabila kredit tersebut macet) (Menarianti, 2015). Namun, perlu diperhatikan juga bahwa debitur yang telah disetujui juga tidak semuanya pembayar kredit yang baik, artinya ada beberapa debitur yang telah disetujui tapi beberapa bulan kemudian pembayarannya menunggak. Oleh karena itu Sistem Pendukung Keputusan salah satu komponen yang cukup penting dalam sistem informasi. Dengan permasalahan tersebut, maka perlu adanya solusi pemecahan masalah yang ada, dengan membuat suatu Sistem Pendukung Keputusan. Dengan demikian Sistem Pendukung Keputusan ini nantinya dapat memenuhi harapan yang ingin dicapai. Namun, perlu diperhatikan juga bahwa nasabah yang telah disetujui juga tidak semuanya pembayar kredit yang baik, artinya ada beberapa nasabah yang telah disetujui tapi beberapa bulan kemudian pembayarannya lebih dari batas jatuh tempo atau bahkan menunggak. Pembayaran yang tidak tepat waktu jika tidak diwaspadai sejak dini maka akan menjadi suatu faktor kerugian bagi perusahaan tersebut.

Ada beberapa penelitian dan tehnik analisa kredit yang dibuat oleh beberapa peneliti seperti: (Jiang, 2009) membuat model untuk memprediksi nasabah yang bermasalah dan tidak bermasalah dalam pembayaran kredit dengan menggunakan model algoritma C4.5. Data yang digunakan diambil dari perusahaan German credit yang merupakan perusahaan pembiayaan. Jiang mengambil beberapa atribut dan kemudian dimasukkan ke dalam model untuk memprediksi persentase nasabah yang bermasalah. Pada penelitian ini, peneliti menyatakan hasil penelitiannya yaitu Statistik menunjukkan bahwa biaya misclassifying kredit lancar dan kredit macet adalah 5 ~ 20 kali dari misclassifying kredit lancar dan kredit macet. (Menarianti, 2015) Membuat model komparasi dengan menggunakan cross validation, confusion matrix, ROC curve dan T-Test pada beberapa algoritma klasifikasi data mining pada kasus nasabah kredit mendapatkan hasil algoritma yang paling akurat adalah algoritma Logistic Regression. karena memiliki nilai akurasi tertinggi yaitu 87,41% dengan uji T-test paling dominan terhadap algoritma lainnya, dengan nilai AUC paling tinggi yaitu 1.000. Algoritma Neural Network walaupun nilai AUC-nya kecil yaitu 0.565 tetapi setelah dilakukan uji T-Test, algoritma ini memiliki sifat yang dominan dengan nilai akurasi cukup tinggi yaitu 86,73% sehingga dapat disimpulkan bahwa algoritma ini cukup akurat. Algoritma Discriminant Analysis dan Decision Tree meskipun memiliki akurasi dan nilai AUC yang tinggi yaitu 87,41 % dan 1.000, tetapi berdasarkan uji T-Test bukan merupakan algoritma yang dominan namun masih cukup baik untuk kasus penentuan pemberian kredit. Algoritma yang memiliki kinerja kurang memuaskan adalah Support Vector Machine, meskipun nilai akurasinya sebesar 86,39% dan Naive Bayes dengan tingkat akurasi sebesar 83,56%. Sedangkan K-Nearest Neighbor merupakan algoritma dengan nilai paling rendah yaitu dengan tingkat akurasi sebesar 76,71%. Dengan kata lain seleksi fitur yang telah dilakukan mempengaruhi hasil akurasi. Tingkat akurasi yang dicapai dapat membantu para analis kredit dalam pengambilan keputusan mengenai pemberian kredit bagi nasabah koperasi. (Kurniawan & Kriestanto, 2016) dengan menggunakan teknologi di bidang data mining yang

mengoptimasi proses pencarian informasi prediksi dalam basis data yang besar, serta menemukan pola-pola yang tidak diketahui sebelumnya. Naïve Bayes memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya dengan mempelajari korelasi hipotesis. Penelitian yang merupakan label kelas yang menjadi target pemetaan dalam klasifikasi dan evidence yang merupakan fitur-fitur yang menjadi masukan dalam model klasifikasi. Pengolahan data berbasis data mining tersebut diharapkan dapat digunakan sebagai alat bantu dalam memprediksikan kelayakan kredit yang diperkirakan layak atau tidaknya pemohon atau nasabah untuk diberikan kredit dengan tingkat akurasi 80,56%. Dan penelitian yang dilakukan oleh (Firmansyah, 2011), juga melakukan penelitian dengan judul “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi” Nilai accuracy, precision, dan recall nya dari data training dapat dihitung dengan menggunakan Rapid Miner 5. Setelah diuji coba dengan metode cross validation, didapatkan hasil pengukuran terhadap data trainingnya yaitu hanya mencapai accuracy = 79.50%, precision = 86.50% dan recall = 91.00% [5]. Hasil pengujian tersebut berdasarkan 5 parameter, merupakan masalah penelitian Teknik Informatika yang masih bisa di tingkatkan akurasinya. Mengapa Hasil penelitian sebelumnya “Penerapan Algoritma Klasifikasi C4.5 untuk Penentuan Kelayakan Pemberian Kredit Koperasi” Nilai accuracy nya hanya 79.50 %? oleh karena itu penulis pada kesempatan ini untuk meningkatkan nilai accuracy dengan memilih model algoritma Boosted Gradient Trees dengan melakukan menggunakan data sampel yang sama.

### **Metode**

Desain riset yang digunakan dalam penelitian ini adalah metode eksperimen. Metode ini menguji kebenaran sebuah hipotesis dengan statistik dan menghubungkannya dengan masalah penelitian (Sugiyono, 2012). Desain eksperimen juga dibagi dua, yaitu eksperimen absolut dan eksperimen komparatif. Eksperimen absolut mengarah kepada dampak yang dihasilkan dari eksperimen. Sedangkan eksperimen komparatif yaitu membandingkan dua objek yang berbeda, misalnya membandingkan dua algoritma yang berbeda dengan melihat hasil statistik masing-masing mana yang lebih baik (Sugiyono, 2012). Dalam penelitian ini, penulis mengambil model eksperimen absolut.

Ada dua tipe dalam pengumpulan data, yaitu pengumpulan data primer dan pengumpulan data sekunder. Data primer adalah data yang dikumpulkan pertama kali dan untuk melihat apa yang sesungguhnya terjadi. Data sekunder adalah data yang sebelumnya pernah dibuat oleh seseorang baik di terbitkan atau tidak (Sugiyono, 2012). Dalam pengumpulan data primer, penulis menggunakan metode observasi dan interview, dengan menggunakan data-data yang berhubungan pengajuan kredit dan bertanya secara langsung kepada pihak yang terlibat secara langsung di dalam sistem. Sedangkan dalam pengumpulan data sekunder menggunakan buku, jurnal, publikasi dan lain-lain

### **Hasil dan Pembahasan**

Algoritma Gradient Tree Boosting menggunakan decision trees dari ukuran yang tetap sebagai weak learner. Decision trees mempunyai sejumlah kemampuan yang membuatnya bernilai untuk boosting, yaitu kemampuan untuk menangani jenis data campuran dan kemampuan untuk memodelkan fungsi yang kompleks.

Untuk mendapatkan nilai kedalaman pohon (max depth) dalam penelitian ini ditentukan dengan cara melakukan uji coba memasukkan nilai dengan range 1 sampai dengan 20 untuk max depth, serta nilai 0.1 untuk learning rate dan 1 untuk sample rate.

Berikut ini adalah hasil dari percobaan yang telah dilakukan untuk penentuan nilai kedalaman pohon (max dept):

Tabel 4.1

Eksperimen penentuan nilai kedalaman pohon (max dept)

| NO           | ACCURACY | AUC   | NUM OF TREES | MAX DEPTH | LEARNING RATE | SAMPLE RATE |
|--------------|----------|-------|--------------|-----------|---------------|-------------|
| Percobaan 1  | 78.66%   | 0.702 | 20           | 1         | 0.1           | 1           |
| Percobaan 2  | 77.41%   | 0.701 | 20           | 2         | 0.1           | 1           |
| Percobaan 3  | 80.33%   | 0.64  | 20           | 3         | 0.1           | 1           |
| Percobaan 4  | 79.91%   | 0.666 | 20           | 4         | 0.1           | 1           |
| Percobaan 5  | 80.74%   | 0.686 | 20           | 5         | 0.1           | 1           |
| Percobaan 6  | 80.74%   | 0.686 | 20           | 6         | 0.1           | 1           |
| Percobaan 7  | 81.58%   | 0.678 | 20           | 7         | 0.1           | 1           |
| Percobaan 8  | 81.16%   | 0.671 | 20           | 8         | 0.1           | 1           |
| Percobaan 9  | 81.16%   | 0.677 | 20           | 9         | 0.1           | 1           |
| Percobaan 10 | 81.16%   | 0.677 | 20           | 10        | 0.1           | 1           |
| Percobaan 11 | 81.16%   | 0.677 | 20           | 11        | 0.1           | 1           |
| ...          | ...      | ...   | ...          | ...       | ...           | ...         |
| Percobaan 20 | 81.16%   | 0.677 | 20           | 20        | 0.1           | 1           |

Nilai learning rate dipilih berdasarkan nilai accuracy terbesar yang dihasilkan. Berdasarkan hasil percobaan di atas, dipilih nilai learning rate sebesar 0.7. Nilai 0.7 ini selanjutnya dipakai untuk percobaan dalam menentukan nilai sample rate.

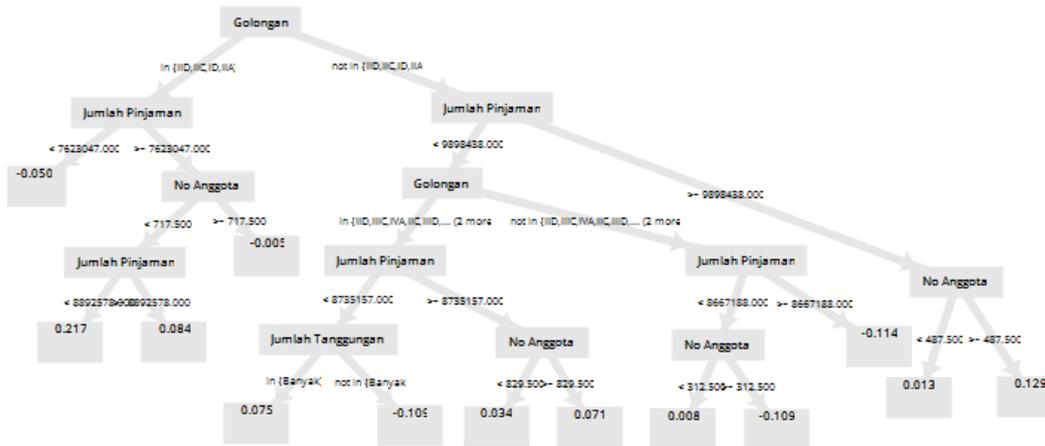
Nilai sample rate ditentukan dengan cara melakukan dengan uji coba memasukkan nilai dengan range 0 sampai dengan 0.9. Nilai max depth 7 dan learning rate 0.7 dipilih berdasarkan percobaan sebelumnya. Berikut ini adalah hasil dari percobaan yang telah dilakukan untuk penentuan nilai sample rate:

Tabel 4.3  
Eksperimen penentuan nilai sample rate

| NO           | ACCURACY | AUC   | NUM OF TREES | MAX DEPTH | LEARNING RATE | SAMPLE RATE |
|--------------|----------|-------|--------------|-----------|---------------|-------------|
| Percobaan 1  | 76.18%   | 0.768 | 20           | 7         | 0.7           | 0.1         |
| Percobaan 2  | 74.55%   | 0.628 | 20           | 7         | 0.7           | 0.2         |
| Percobaan 3  | 69.95%   | 0.623 | 20           | 7         | 0.7           | 0.3         |
| Percobaan 4  | 76.21%   | 0.675 | 20           | 7         | 0.7           | 0.4         |
| Percobaan 5  | 81.20%   | 0.637 | 20           | 7         | 0.7           | 0.5         |
| Percobaan 6  | 79.96%   | 0.631 | 20           | 7         | 0.7           | 0.6         |
| Percobaan 7  | 83.30%   | 0.675 | 20           | 7         | 0.7           | 0.7         |
| Percobaan 8  | 84.55%   | 0.627 | 20           | 7         | 0.7           | 0.8         |
| Percobaan 9  | 85.36%   | 0.606 | 20           | 7         | 0.7           | 0.9         |
| Percobaan 10 | 86.63%   | 0.59  | 20           | 7         | 0.7           | 1           |

Berdasarkan hasil percobaan di atas, maka untuk parameter gradient boosted trees dipilih nilai 7 untuk max depth, 0.7 untuk learning rate dan 1 untuk sample rate.

Dari eksperimen terbaik di atas maka didapat model tree dengan menghasilkan tujuh kedalaman pohon dengan model tree yang terbentuk pada gambar 4.1



Gambar 4.1

Gambar Tree dengan Max depth 7 dengan model Gradient Boosted Trees

## Kesimpulan

Dari hasil penelitian terbukti bahwa algoritma Gradient Boosted Trees lebih akurat dibandingkan analisa yang dilakukan oleh analis kredit. Hal ini dibuktikan dengan hasil evaluasi penelitian bahwa algoritma Gradient Boosted Trees mampu menganalisa kredit yang bermasalah dan yang debitur yang tidak bermasalah sebanyak 86.63%.

## Daftar Pustaka

- Firmansyah. (2011). Penerapan Algoritma Klasifikasi C4.5 Untuk Penentuan Kelayakan Pemberian Kredit Koperasi. Jakarta: Stmik Nusa Mandiri.
- Jiang, Y. (2009). Credit Scoring Model Based On Decision Tree And The Simulated Annealing Algorithm. World Congress On Computer Science And Information Engineering (Pp. 18 - 22). Los Angeles: Ieee Computer Society.
- Kotsiantis, S., Kanellopoulos, D., Karioti, V., & Tampakas, V. (2009). An Ontology-Based Portal For Credit Risk Analysis. Ieee International Conference On Computer Science And Information Technology, (Pp. 165-169). Beijing.
- Kurniawan, D. A., & Kriestanto, D. (2016). Penerapan Naïve Bayes Untuk Prediksi Kelayakan Kredit . Jurnal Informatika Dan Komputer (Jiko), 19 - 23.
- Lai, K. K., Lean Yu , L., & Zhou, L. (2006). Credit Risk Evaluation With Least Square Support Vector Machine. Springer-Verlag , 490-495.
- Larose, D. (2005). Discovering Knowledge In Data. New Jersey: John Willey & Sons, Inc.
- Menarianti, I. (2015). Klasifikasi Data Mining Dalam Menentukan Pemberian Kredit Bagi Nasabah Koperasi. Jurnal Ilmiah Teknosains, 36-45.
- Odeh, O., Featherstone, A., & Das, S. (2010). Predicting Credit Default: Comparative Results From An Artificial Neural Network, Logistic Regression And Adaptive Neuro-Fuzzy Inference System. Eurojournals Publishing, Inc, 7 - 17.
- Sucipto, A. (2015). Prediksi Kredit Macet Melalui Perilaku Nasabah Pada Koperasi Simpan Pinjam Dengan Menggunakan Metode Algoritma Klasifikasi C4.5. Disprotek, 75 - 87.

- Sugiyono. (2012). *Metode Penelitian Kuantitatif Kualitatif Dan R&D*. Bandung: Alfabeta.
- Sunjana. (2010). Klasifikasi Data Nasabah Sebuah Asuransi Menggunakan Algoritma C4.5. *Seminar Nasional Aplikasi Teknologi Informasi 2010 (Snati 2010)* , 31 - 34.
- Susanto, Uryani, N. S., & Hidayat, C. R. (2014). Penerapan Algoritma C4.5 Untuk Menganalisis Kelayakan Pemberian Kredit Nasabah. *Seminar Nasional*.
- Triwahyuniati, N. (2008). Pelaksanaan Analisis Kredit Pada di PT Bank Huga Semarang. *Thesis Magister Kenotariatan* , 38.
- Witten, Frank, & Hall. (2011). *Data Mining : Practical Machine Learning and Tools*. Burlington: Morgan Kaufmann Publisher.
- Zurada, J. (2010). Could Decision Trees Improve the Classification Accuracy and Interpretability of Loan Granting Decisions. *HICSS '10 Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*. (p. 19). Hawaii: Koloa.