

PEMILIHAN FITUR KEPUTUSAN KREDIT BERBASIS *MAXIMAL INFORMATION COEFFICIENT* (MIC) : SUATU GAGASAN AWAL

Hernawati Pramesti

Fakultas Ekonomi Universitas Kristen Surakarta

e-mail: hernawati73@yahoo.co.id

ABSTRACT

The need for effective risk management means that banks must begin to look for continuous improvement in the techniques used for credit analysis by producing the development and application of various quantitative models. - payments from consumers and the risk of default. This initial idea will be carried out using an approach that has never been done in studies related to credit scoring, where the method of selecting the independent variable Maximal Information Coefficient (MIC) to get the best credit decision.

Key word: Feature selection, Maximal Information Coefficient, Credit decision

PENDAHULUAN

Perbankan sebagai lembaga jasa keuangan harus mampu memberikan pelayanan yang memuaskan nasabahnya. Upaya ini berguna untuk meningkatkan hubungan yang harmonis antara perbankan dengan nasabahnya sehingga dapat mendorong terciptanya loyalitas nasabah yang nantinya akan memberikan rekomendasi kepada nasabah atau calon peminjam lainnya yang menguntungkan bagi perbankan. Kondisi ini akan berdampak pada meningkatnya jumlah pengguna produk perbankan tersebut (Xia et al., 2017) (Dendawijaya, 2003; Saptono dan Widiyatmanta, 2006; Bramantya dan Ronny, 2007).

Pengelolaan kredit merupakan salah satu produk perbankan, dengan dilakukannya pengelolaan kredit secara profesional diharapkan dapat meningkatkan kualitas aset, likuiditas dan profitabilitas bank, disebabkan kualitas aset, tingkat likuiditas dan profitabilitas yang tinggi menunjukkan kinerja perbankan yang tinggi pula. Oleh sebab itu diperlukan suatu manajemen kredit yang menjadi pengelolaan kredit yang baik yang dimulai dari perencanaan jumlah kredit, penentuan suku bunga kredit, prosedur pemberian kredit, analisis pemberian kredit sampai pengendalian dan pengawasan kredit macet (Kasmir, 2001; Harahap, 2009). Salah satu penyebab kredit bermasalah (resiko kredit) berasal dari pihak internal, yaitu kurang telitnya tim dalam melakukan survei dan analisis atau bisa juga karena penilaian dan analisis yang bersifat subjektif. Penyebab ini dapat diatasi dengan sistem komputer (Amseke and Winarko, 2014), namun menurut (Kasmir, 2001)

Berdasarkan penelitian-penelitian yang pernah dilakukan, evaluasi resiko kredit merupakan masalah yang menarik dalam analisa keuangan (Leidiyana, 2013; Chaidir, 2014; Xia et al., 2017). Penelitian mengenai analisis kelayakan pemberian kredit untuk konsumen khususnya kredit kepemilikan barang dengan metode klasifikasi data mining telah banyak dilakukan (Angelini et al., 2008; Oreski et al., 2012; Chavira et al., 2017; Chi and Li, 2017)(Yu et al., 2008; Karlan and Zinman, 2009; Twala, 2010; Neto et al., 2017) ada juga yang menggunakan metode klasifikasi data set yang berhubungan dengan resiko kredit (Karlan and Zinman, 2009; Twala, 2010; Neto et al., 2017). Dalam pembahasan tentang resiko

kredit banyak penelitian yang menyoroti tentang metode atau sistem yang digunakan untuk penentuan resiko kredit dan terdapat beberapa penelitian yang terkait dengan penerapan algoritma genetika untuk penentuan resiko kredit (Oreski and Oreski, n.d.; Twala, 2010; Oreski et al., 2012; Leidiyana, 2013; Chavira et al., 2017; Metawa et al., 2017; Neto et al., 2017; Zhang et al., 2018). Sejalan dengan pertumbuhan bisnis, kredit memang merupakan masalah yang menarik untuk diteliti, selain menggunakan Algoritma Genetika terdapat beberapa riset bidang komputer yang terkait dengan cara untuk mengurangi resiko kredit telah banyak dilakukan dalam rangka menghindari kehancuran suatu perusahaan pembiayaan (Leidiyana, 2013; Chaidir, 2014; Accornero et al., 2017). Dengan kata lain, klasifikasi yang dikembangkan dengan data kredit guna membangun sistem pendukung keputusan, membantu bank dalam memutuskan apakah memberikan pinjaman atau tidak. Menurut Xia (2017) model prediktif yang diterapkan pada *credit scoring* dapat dibagi menjadi dua kelompok: pendekatan statistik dan metode kecerdasan buatan (*Artificial Intelligence / AI*). Sejumlah penelitian telah terfokus pada metode pemodelan yang menawarkan algoritma baru untuk meningkatkan akurasi penilaian kredit, metode ini meliputi metode statistik seperti *Linear Discriminate Analysis* (LDA) (Altman, 1968; Lee et al., 2002) dan *Logistic Regresion* (LR) (Wiginton, 1980; Leidiyana, 2013; Figini and Giudici, 2017) dan metode kecerdasan buatan, seperti artificial Neural Network (NN) (Trippi and Turban, 1992; Angelini et al., 2008; Yu et al., 2008), mesin vektor port (Min and Lee, 2005; Huang et al., 2007) dan Decision Tree (DT) (Nie et al., 2011; Xia et al., 2017).

Sejauh ini penelitian mengenai resiko kredit yang dikaitkan dengan *credit scoring* telah banyak dilakukan dan penelitian tersebut belum ada yang menggunakan metode *Maximal Information Coefficient* (MIC). MIC merupakan metode pengukuran dependensi antar 2 variabel acak yang dikembangkan oleh Reshef dkk., 2011, yang diharapkan dapat digunakan untuk memperoleh variabel independen yang relevan terhadap keputusan pemberian kredit. Berdasarkan rangking, MIC menghasilkan variabel independen yang paling berpengaruh terhadap keputusan pemberian kredit.

Model sistem menggunakan metode MIC ini diharapkan dapat memaksimalkan objektivitas penilaian terhadap variabel independen yang digunakan dalam proses penentuan keputusan kredit, mengurangi waktu komputasi yang dibutuhkan untuk proses penentuan keputusan kredit, serta menghasilkan hasil keputusan kredit yang optimal dan akurat. Model yang dihasilkan diharapkan dapat mengurangi biaya yang harus dikeluarkan dalam melakukan proses keputusan kredit, yang selanjutnya akan meningkatkan *income* kepada pihak pemberi kredit dalam hal ini perbankan. Penggunaan teknologi dapat membantu mengurangi biaya tenaga kerja terutama biaya tenaga kerja yang berhubungan langsung dengan pengambilan keputusan akan berkurang, penggunaan teknologi yang tepat juga bisa membantu kegiatan organisasi lebih baik dan menyediakan banyak lagi pelayanan yang konsisten untuk pelanggan, dalam hal ini penelitian telah menunjukkan bahwa peningkatan persepsi layanan dan kualitas produk mengarah ke peningkatan kepuasan pelanggan dan laba (Karat, Zahorik, dan Keiningham, 1998; Kimes, 2008). Teknologi juga dapat menjadi tambahan yang berharga untuk mencapai profitabilitas yang lebih besar, sebab jika teknologi diimplementasikan dengan benar teknologi akan

meningkatkan persepsi kontrol pelanggan dan kenyamanan dan dengan demikian meningkatkan kepuasan pelanggan, menambah prospek peningkatan bisnis berulang, maka dapat disimpulkan bahwa penggunaan teknologi dapat membantu meningkatkan kecepatan layanan (yang akan menghasilkan peningkatan volume selama periode sibuk), mengurangi biaya tenaga kerja, dan menarik tambahan pendapatan (Kimes, 2008)

Penelitian yang dilakukan oleh (Louzada et al., 2016) dengan mereview 187 paper menyatakan bahwa terdapat 9 teknik klasifikasi Credit Scoring yaitu *Neural Network, Support Vector Machine, Linear Regression, Decision Tree, Logistic Regression, Fuzzy Logic, genetic programming, discriminant analysis dan bayesian network*. Sedangkan penelitian dengan menggunakan metode MIC tidak termasuk dalam 9 teknik klasifikasi tersebut, maka penelitian dengan menggunakan metode ini diharapkan dapat membantu analisis kredit dalam mengambil keputusan terkait pemberian kredit terhadap calon penerima kredit dalam hal: mengurangi tingkat subjektivitas analisis kredit dalam keputusan kredit dan mengurangi waktu proses keputusan kredit, karena selain menghasilkan korelasi antara variabel independen dan dependen juga menghasilkan sekelompok fitur yang paling relevan, sehingga dengan kelompok fitur yang paling relevan yang digunakan untuk keputusan kredit maka akan diperoleh keputusan kredit dalam waktu yang relative singkat daripada menggunakan seluruh fitur. Dengan berkurangnya waktu serta personal pada bagian keputusan kredit maka akan membuat perbankan memperoleh tambahan pendapatan yang selanjutnya juga akan meningkatkan kinerja keuangan perbankan yang bersangkutan. Berdasarkan hal tersebut maka muncul pertanyaan, bagaimana penggunaan metode MIC dapat membantu analisis kredit dalam memberikan keputusan kredit ?

KAJIAN LITERATURE

1. KEPUTUSAN KREDIT

Ketentuan yang melandasi kegiatan kredit salah satunya adalah Undang—Undang Perbankan No 10 Tahun 1992 menyebutkan bahwa : kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan peminjam untuk melunasi hutangnya setelah jangka waktu tertentu dengan sejumlah bunga, imbalan atau pembagian hasil keuntungan. Demikian juga menurut Claudia (2014) Kredit adalah cara menjual barang dan atau pinjaman uang dengan pembayaran secara tidak tunai dimana pembayaran ditangguhkan atau diangsur dengan pinjaman sampai batas jumlah tertentu yang diizinkan oleh bank atau badan lain.

Keputusan kredit yang dilembagakan dan penilaian perorangan dari petugas atau pejabat pinjaman adalah penggerak utama untuk keputusan kredit, Rodgers (1999) mengemukakan bahwa keputusan kredit juga sangat dipengaruhi oleh faktor kontinjensi, termasuk prospek perdagangan, ukuran perbankan dan budaya perbankan. Literatur tentang keputusan kredit sebagian besar ditandai oleh beragamnya deskripsi bagaimana prosesnya dapat ditingkatkan. Dua dimensi proses kredit dianggap sangat bermasalah: (1) proses analisis informasi dan (2) sumber informasi. Sehubungan dengan dimensi pertama mendasarkan pada proses psikologis yang menghasilkan tingkah laku. Mengenai dimensi kedua, yaitu sumber informasi terdapat banyak kemungkinan dan biasanya, laporan tahunan

digunakan untuk membantu mengungkapkan situasi keuangan perusahaan (Rodgers, 1991). Sedangkan definisi keputusan kredit adalah keputusan peminjam tentang jumlah kredit, tingkat suku bunga dan jumlah agunan (Catusus, 2003). Yang termasuk dalam keputusan kredit adalah memperpanjang kredit, jumlah pinjaman, jumlah pelunasan yang dijadwalkan, jangka waktu pinjaman untuk memaksimalkan return, dan keputusan kredit juga didasarkan pada keuntungan yang diharapkan serta resiko kredit untuk jangka waktu kontrak (Tanaka, 2003).

Salah satu penyebab kredit bermasalah berasal dari pihak internal, yaitu kurang telitinya tim dalam melakukan survei dan analisis atau bisa juga karena penilaian dan analisis yang bersifat subjektif (Amseke and Winarko, 2014). Pada umumnya pihak bank dalam mengadakan penilaian terhadap kredit harus memperhatikan beberapa persyaratan. Adapun persyaratan tersebut dikenal dengan prinsip 5C (Anwar, 2013; Chaidir, 2014; Lailiyah, 2014), Ikasari (2014) serta Cronje dan Atahau (2017) yaitu : a). *Character*, karakter menunjukkan kemungkinan atau probabilitas dari langganan untuk secara jujur berusaha untuk memenuhi kewajiban-kewajibannya. b). *Capacity*, kapasitas dapat diukur dari catatan dimasa yang lampau, dilengkapi dengan observasi fisik pada pabrik atau toko langganan. c). *Capital*, faktor modal dapat diukur dari posisi keuangan perusahaan secara umum, dimana hal ini ditunjukkan oleh beberapa analisa ratio keuangan yang khususnya ditekankan pada aktiva yang dimiliki sendiri oleh perusahaan. d). *Collateral*, Faktor jaminan dicerminkan oleh aktiva dan langganan yang dikaitkan atau dijaminan lagi keamanan kredit yang diberikan kepada langganan tersebut dan e). *Conditions*, kondisi ekonomi menunjukkan pengaruh langsung dari trend ekonomi pada umumnya terhadap perusahaan yang bersangkutan

2. METODE CREDIT SCORING

Metode *credit scoring* digunakan untuk mengklasifikasikan konsumen yang mengajukan kredit termasuk dalam kategori baik atau buruk dan *credit scoring* juga mencoba untuk mengelompokkan keragaman karakteristik calon peminjam berdasarkan baik ketaatan, kesalahan ataupun kesalahan dalam memenuhi kewajibannya, sehingga fungsi *credit scoring* selain membantu bank menentukan apakah pinjaman disetujui atau tidak juga menentukan berapa besarnya pinjaman yang akan diperoleh, berapa *term* yang didapatkan serta berapa besarnya bunga pinjaman yang dikenakan (Amarnath, 2004; Dendy, 2011).

2.1. Kajian Literature Terkait Metode Pemilihan Fitur untuk Credit Scoring

Secara umum, pemilihan fitur dapat dikategorikan kedalam 2 jenis: filter dan wrapper (Rodriguez, 2010). Metode filter memilih fitur terbaik dengan mengevaluasi sifat dasar data, sehingga cepat dan mudah diterapkan. Metode Wrapper memilih fitur terbaik sesuai dengan akurasi classifier, membuat hasilnya cocok dengan algoritma klasifikasi yang telah ditentukan. Memilih antara metode filter dan wrapper bukanlah tugas yang mudah, karena tergantung pada banyak faktor. Untuk tujuan umum penghapusan fitur tidak relevan, filter adalah pilihan yang baik karena tidak bias dan cepat. Di sisi lain, untuk meningkatkan kinerja klasifikasi, wrapper lebih cocok karena lebih sesuai untuk tugas klasifikasi.

Kajian literature terkait metode pemilihan fitur untuk *credit scoring*

Peneliti & Tahun	Judul	Pemasalahan	Metode	Hasil	Keterbatasan/Pe nelitian lanjutan
Y Liu dan M Schumann, 2005	Data mining feature selection for credit scoring models	Fitur yang digunakan memiliki pengaruh penting terhadap performansi model credit scoring. Proses memilih fitur terbaik untuk model biasanya bersifat tidak sistimatis dan didominasi oleh percobaan yang semauanya.	<ol style="list-style-type: none"> Menggunakan 4 metode pemilihan fitur ('Relief', 'Correlation-based', 'Consistency-based' dan 'Wrapper') untuk mereduksi fitur dan meningkatkan performansi model scoring yaitu: <i>model simplicity, model speed and model accuracy</i>. Eksperimen dilakukan terhadap real data sets (German credit insurance company) menggunakan 4 algoritma klasifikasi : model tree (M5), neural network (multi-layer perceptron with back-propagation), logistic regression, dan k-nearest-neighbours. 	<ol style="list-style-type: none"> Metode pemilihan fitur berbasis <i>consistency</i> dan <i>wrapper</i> memiliki kinerja yang relatif lebih baik. Setelah pemilihan fitur, peningkatan akurasi model ditampilkan algoritma k-NN, tapi tidak ada perbaikan ditunjukkan untuk algoritme lain. Pengurangan jumlah fitur mengurangi waktu pelatihan dan menyederhanakan model terakhir. Kadang dalam prakteknya, perlu dan layak mengurangi ruang fitur walaupun perlu mengorbankan akurasi, khususnya saat memakai sejumlah besar fitur. 	<ol style="list-style-type: none"> Perlu dilakukan eksperimen menggunakan dataset, metode pemilihan fitur, dan metode klasifikasi yang berbeda sebagai perbandingan. Sangat berarti untuk mengetahui dampak pemilihan fitur yang digunakan terhadap kriteria klasifikasi yang lain.
Jian SHI, Shu-you ZHANG, Le-miao QIU, 2013	Credit scoring by feature-weighted support vector machines	Krisis hutang & keuangan dewasa ini menjadikan credit risk management sebagai isu penting dalam financial research. Model credit scoring yang handal sangat penting bagi lembaga keuangan untuk mengevaluasi aplikasi kredit dan banyak dipelajari di bidang pembelajaran mesin dan statistik	Feature-weighted support vector machine (SVM) digunakan untuk menilai credit risk dimana <i>F-score</i> is diadopsi untuk meranking fitur penting. Dengan mempertimbangkan interaksi bersama antar fitur, Random Forest digunakan untuk mengukur (secara relatif) fitur penting pada datasets (German and Australian credit datasets from the UCI Repository of Machine Learning Databases)	Dibandingkan dengan SVM tradisional, fitur-weighted SVM memberikan fitur input dengan derajat kepentingan berbeda untuk meningkatkan hyperplane terpisah dalam ruang fitur dimensi tinggi yang ditransformasikan. Dua strategi pembobotan fitur, F-score dan Random Forest , diadopsi untuk mengukur fitur penting. Hasil eksperimen menunjukkan random forest feature-weighted SVM (RFFWSVM) memberikan akurasi klasifikasi tinggi dan stabilitas baik untuk aplikasi credit scoring praktis.	
Waad Bouaguel, Ghazi Bel Mufti, and Mohamed Limam, 2013	Rank Aggregation for Filter Feature Selection in Credit Scoring	<ol style="list-style-type: none"> Industri kredit tumbuh cepat, lembaga kredit mengumpulkan data pelanggan kredit dan digunakan untuk membangun model kredit. Informasi yang dikumpulkan penuh dengan fitur yang tidak diinginkan Metode seleksi fitur Filter mengungguli teknik pemilihan fitur lainnya karena efektif dan cepat komputasi. 	<ol style="list-style-type: none"> Mengusulkan model fusi pemilihan fitur menggabungkan hasil dengan metode pemilihan fitur filter berbeda (relief (Euclidean distance), Pearson correlation coefficient (PCC), Mutual Information (MI), majority vote, dan Mean Aggregation) melalui teknik agregasi. 4 dataset (Australian and German credit datasets) and a dataset from a Tunisian bank and the HMEQ dataset) digunakan menguji performansi metode 3 Metode klasifikasi (Logistic Regression (LR), Naive Bayes (NB) and Support Vector Machine (SVM)) digunakan untuk menguji performansi model. 	Hasil menunjukkan bahwa ada efek menguntungkan secara umum dari penggabungan peringkat fitur dibandingkan dengan yang dihasilkan oleh metode tunggal. Bahkan kinerja fusi lebih baik atau setidaknya sedekat metode filter.	Sebagai tambahan untuk pekerjaan ini, memilih jumlah fitur yang tepat adalah sebuah tantangan, namun untuk memilih jumlah fitur yang sesuai dari daftar peringkat masih merupakan masalah terbuka untuk dipelajari.
Van-Sang Ha, Ha-Nam Nguyen, 2016	Credit scoring with a feature selection approach based deep learning	Model penilaian kredit yang andal sangat penting bagi lembaga keuangan untuk mengevaluasi aplikasi kredit dan telah dipelajari secara luas di bidang pembelajaran mesin dan statistik.	<ol style="list-style-type: none"> Menggunakan 2 public datasets: Australia and German credit untuk menguji metode yang diusulkan. Mengusulkan metode seleksi fitur berdasarkan yang terbaik dari Rata-rata + Median Skor dan standar deviasi terendah (SD), kemudian diintegrasikan dengan pengklasifikasi deep learning untuk mengestimasi performansi: cross validation accuracy dan penting/manfaat setiap fitur dalam proses training data set. 	<ol style="list-style-type: none"> jumlah fitur sedikit memungkinkan departemen kredit berkonsentrasi mengumpulkan variabel yang relevan dan penting. Prosedur pemrosesan paralel menyebabkan penurunan signifikan dalam waktu proses. Sehingga, beban kerja personil kredit kurang. Hasil eksperimen menunjukkan metode yang diusulkan efektif. Evaluasi dilakukan lebih cepat dan meningkatkan akurasi. 	
Van-Sang Ha and Ha-Nam	FRFE: Fast Recursive Feature Elimination	Pembangunan model credit scoring dewasa ini menjadi topik hangat dan umumnya menggunakan	1. Mengusulkan penggunaan H2O parallel random forest (PRF) untuk mengestimasi performansi dan mereduksi running time.	<ol style="list-style-type: none"> Akurasi classifier menggunakan fitur terpilih lebih baik n. Dengan lebih sedikit fitur, memungkinkan departemen kredit 	

Peneliti & Tahun	Judul	Pemasalahan	Metode	Hasil	Keterbatasan/Pe nelitian lanjutan
Nguyen, 2016	for Credit Scoring	teknik data mining. Tidak semua fitur dapat digunakan untuk proses klasifikasi, karena adanya fitur yang Irrelevant and redundant dalam data kredit, yang akan berpengaruh pada berkurangnya akurasi klasifikasi.	2. Pada fase pertama, data training set was dilatih dan diuji menggunakan PRF untuk mendapatkan fitur terbaik. Pada fase kedua, hasil fase pembelajaran (learning phase) digunakan sebagai filter untuk menguji dataset (Australia and German credit).	konsentrasi di pengumpulan variabel esensial dan relevan. 3. Pemrosesan paralel menghasilkan pengurangan waktu proses. Beban kerja personil kredit jadi berkurang, mereka tidak perlu memperhitungkan sejumlah besar fitur saat melakukan evaluasi. 4. Metode yang diusulkan efektif menganalisis resiko kredit.	
Damodar Reddy Edla · Diwakar Tripathi · Ramalinga swamy Cheruku · Venkatana reshababu Kuppili, 2017	An Efficient Multi-layer Ensemble Framework with BPSOGSA-Based Feature Selection for Credit Scoring Data Analysis	Credit scoring banyak digunakan oleh industri kredit dan lembaga keuangan untuk pengambilan keputusan keuangan. Ini adalah cara untuk menilai risiko yang terkait dengan pemohon berdasarkan data historis. Namun, data historis mungkin memiliki sejumlah besar fitur redundan dan noise yang dapat mempengaruhi kinerja model credit scoring.	1. Hybrid credit scoring model yang diusulkan hybrid binary particle swarm optimization (BPSOGSA) untuk pemilihan fitur dan multi-layer ensemble classifier framework dengan 5 classifiers yang berbeda (multi-layer feed forward neural network (MLFFNN), Naïve Bayes (NB), quadratic discriminant analysis (QDA), time delay neural network (TDNN), dan deep tensor neural networks (DTNN)). 2. Framework diusulkan diguna kan 4 credit scoring datasets: Australian, German - categorical, German-numerical and Japanese credit scoring datasets .	Hasil eksperimen menunjukkan bahwa fitur terpilih menggunakan fungsi V-shaped transfer dan fungsi fitness yang diusulkan lebih representatif dan dapat meningkatkan kinerja classifier MLFFNN, NB, QDA, DTNN dan TDNN. Secara keseluruhan, untuk semua dataset, model ensemble yang diusulkan mengungguli mode ensemble tradisional, seperti Random Forest (RF), Majority Voting (MV), Layered Majority Voting (LMV), Weighted Voting (WV) dan Layered Weighted Voting (LWV) dalam hal akurasi, sensitifity, dan G-measure. Dapat disimpulkan bahwa ensemble framework yang diusulkan dengan pemilihan fitur berbasis BPSOGSA adalah pendekatan yang efisien untuk credit scoring.	1. Pemilihan fitur dapat diuji pada teknik optimasi Whale and Mouth Flame di fase-1 untuk pemilihan fitur yang efektif. 2. Pendekatan keputusan multi-kriteria untuk penempatan pengklasifikasi dalam fase-2 dapat digunakan sebagai penempatan pengklasifikasi pada layered ensemble framework, karena penempatan pengklasifikasi dapat mempengaruhi kinerja klasifikasi.
ZhanFeng Liu · Su Pan, 2017	Fuzzy-Rough Instance Selection Combined with Effective Classifiers in Credit Scoring	Nomor cluster yang salah atau titik awal yang buruk dari setiap cluster memiliki pengaruh negatif pada akurasi klasifikasi dalam sistem penilaian kredit berbasis hybrid classifier.	usulkan Fuzzy-rough instance selection (FRIS) yang menghilangkan lebih banyak instans daripada yang diperlukan karena penghapusan satu instans dapat mempengaruhi keanggotaan wilayah positif dari instans yang tersisa, sehingga algoritma k-nearest neighbor (KNN) diadopsi untuk memeriksa instans yang dieliminasi. Instans yang tereduksi dimodelkan dengan pengklasifikasi efektif (SVM) dan akurasi terbaik dapat diperoleh dalam credit scoring.	hybrid classifier mengeliminasi instans yang tidak representatif secara efektif, dan meningkatkan performansi classifier ketika mengadopsi parameter yang sesuai. Hybrid classifier yang diusulkan juga mengungguli linear discriminant analysis, logistic regression dan neural networks.	mbangan lanjut dapat dilakukan kan peningkatan performansi pengklasifikasi menggunakan dataset lebih sedikit instans tidak representatif secara simultan menghilangkan instans dan fitur.
Hazar Altinbas · Goktug Cenk Akkaya, 2017	Improving the performance of statistical learning methods with a combined meta-heuristic for consumer credit risk assessment	Dalam aplikasi credit risk, penggunaan statistical learning methods dapat meningkatkan akurasi dari beberapa metode tersebut dengan mengidentifikasi subset fitur optimal dari keseluruhan variabelnya.	German credit risk dataset sebuah benchmark set digunakan memutuskan variabel penjelas yang digunakan dan menentukan peminjam gagal atau tidak. Analisis dilakukan 2 tahap : pertama, menerapkan beberapa pembelajaran statistik klasik (Logistic regression (LR), Linear discriminant analysis (LDA), K-nearest neighbors (KNN), Support vector machines Maximal margin classifiers (MMCs), Random forests (RF), Boosted trees (BTs)). Kedua, hasil gabungan meta-heuristik (Genetic Algorithm dan Simulated Annealing) digunakan metode statistik untuk evaluasi nilai dan melakukan pemilihan fitur secara simultan untuk meningkatkan kinerja prediktor.	Hasil yang menjanjikan dicapai oleh model dan dapat digunakan oleh analis dan praktisi dalam mengawasi otoritas dan lembaga keuangan dalam melakukan penilaian risiko kredit dan pengambilan keputusan.	
Waad	An	Kenyataan, adanya	Sebuah metode ensemble baru untuk seleksi	Pengkombinasian classifiers	peran jumlah

Peneliti & Tahun	Judul	Pemasalahan	Metode	Hasil	Keterbatasan/Pe nelitian lanjutan
Bouaguel and Mohamed Limam, 2015	Ensemble Wrapper Feature Selection for Credit Scoring	perbedaan bias dan asumsi masing-masing classifier dapat mempengaruhi hasil akhir dalam hal akurasi dan waktu eksekusi. Tingkat kerumitan klasifikasi adalah faktor fundamental untuk diselidiki. Secara teori, pengklasifikasi yang kompleks, diperlukan waktu lebih lama untuk memilih bagian fitur terbaik daripada pengklasifikasi yang dianggap sederhana.	fitur diusulkan dan dievaluasi pada empat dataset kredit (Australian dan German credit datasets, Tunisian bank, dan HMEQ dataset). Pendekatan yang diusulkan terdiri dari dua tahap. Pertama, melakukan reduksi dimensi menggunakan pengetahuan ahli bank tentang fitur-fitur redundan yang dihilangkan. Kedua, subset yang dihasilkan dievaluasi menggunakan proses multi-pengklasifikasi dua pendekatan pengaturan, yaitu pendekatan same-type DT, SVM, ANN, dan KNN dan pendekatan mixed-type (J48, Random Forest (RF), Multilayer Perceptron (MP), Voted Perceptron (VP), K=1 (1NN), K=5 (5NN), Polynomial (SVMP), Radial (SVMR))	dapat meningkatkan performansi scoring models. Dari kedua tahap dapat terlihat bahwa penggunaan informasi sebelumnya pada fitur yang relevan secara efektif menginduksi peningkatan yang signifikan dalam kompleksitas dengan peningkatan generalisasi. Selain itu sifat pengklasifikasi mempunyai efek penting pada hasil pemilihan fitur wrapper.	pengklasifikasi dalam pemilihan fitur ensemble perlu diteliti secara mendalam.
Waad Bouaguel and Mohamed Limam, 2016	A New Way for Combining Filter Feature Selection Methods	Rank aggregation menampilkan proses pemilihan fitur dan gabungan hasilnya ke daftar peringkat akhir. Sejumlah metode yang ada menggabungkan hasil individu ke daftar fitur peringkat tunggal	Pendekatan diusulkan terdiri dua tahap: pertama, mengevaluasi kemiripan dan stabilitas metode pemfilteran tunggal (relief, chi-squared, mutual information, and correlation) kedua gabung kan hasil yang stabil. Hasil yang diperoleh pada dataset kredit Australia dan Jerman digunakan SVM dan Decision tree menegaskan peringkat fitur ensemble punya dampak peningkatan performansi.	Hasil pada dua dataset kredit menunjukkan peningkatan yang luar biasa ketika menggunakan metode agregasi peringkat baru yang diusulkan dibandingkan dengan peringkat individual dan metode agregasi kompetitif lainnya yang diambil sebagai input.	akan lebih baik untuk mempelajari ukuran kemiripan lainnya untuk menghitung tingkat stabilitas metode filtering.

2. MIC SEBAGAI METODE CREDIT SCORING YANG DIUSULKAN

2.1. Mutual Information (MI).

Asosiasi antar variabel pada data pinjaman, terhubung secara linear maupun non linear. Untuk itu dibutuhkan pengukuran yang dapat mengakomodir 2 hal tersebut. *Mutual Information* (MI) adalah sebuah pengukuran ketergantungan variabel baik secara linear maupun non linear (Battiti, 1994). MI mengukur 'jumlah informasi' yang diperoleh tentang suatu variabel acak, melalui variabel acak lain. Dalam penelitian ini, MI digunakan untuk mengukur 'informasi' tentang sebuah variabel dependen (resiko kredit), melalui variabel independen.

2.2. Maximal Information Coefficient (MIC)

Maximal Information Coefficient (MIC) adalah metode pengukuran dependensi antar 2 variabel acak yang dikembangkan oleh Reshef (Reshef et al., 2011). MIC menggunakan *binning* (kantong) sebagai sarana untuk menerapkan *Mutual Information* (MI). Pemilihan variabel independen yang relevan berbasis MIC diawali dengan menganalisa hubungan antar variabel, yang bila dilakukan secara manual, cukup kompleks dan sulit, terutama pada dataset yang mengandung jumlah variabel cukup banyak (dataset berdimensi tinggi). Reshef mengeksplorasi dataset berdimensi tinggi ini dengan mencari pasangan variabel yang sangat berhubungan erat. Dimulai dengan mengukur dependensi setiap pasangan variabel, membuat ranking setiap pasangan berdasarkan *score*, dan menguji pasangan yang memiliki *score* tinggi (*top score*).

MIC berbasis pada ide bahwa bila terdapat hubungan 2 variabel, maka sebuah grid dapat digambarkan pada *scatterplot*. Grid mempartisi data sekaligus membungkus (mengandung informasi) hubungan kedua variabel tersebut. Untuk menghitung MIC dari kedua variabel, maka eksplorasi terhadap semua grid perlu dilakukan untuk menghitung informasi bersama (MI) yang mungkin dari setiap

pasangan integer (x,y) yaitu pasangan variabel (independen, dependen). Kemudian melakukan normalisasi nilai MI (dengan maksud memberikan keadilan perbandingan pada semua grid yang berbeda dimensi) dan mendapatkan nilai termodifikasi antara 0 dan 1.

KESIMPULAN

Berdasarkan pembahasan beberapa literature di atas, kami mengusulkan metode MIC sebab dengan menggunakan metode ini diharapkan dapat membantu analis kredit dalam mengambil keputusan kredit dalam hal: mengurangi tingkat subyektivitas analis kredit dalam keputusan kredit dan mengurangi waktu proses keputusan kredit, karena selain menghasilkan korelasi antara variabel independen dan dependen juga menghasilkan sekelompok fitur yang paling relevan, sehingga dengan kelompok fitur yang paling relevan yang digunakan untuk keputusan kredit maka akan diperoleh keputusan kredit dalam waktu yang relative singkat daripada menggunakan seluruh fitur. Dengan berkurangnya waktu serta personil pada bagian keputusan kredit maka akan membuat perbankan memperoleh tambahan pendapatan yang selanjutnya juga akan meningkatkan kinerja keuangan perbankan yang bersangkutan.

DAFTAR PUSTAKA

- Altman, E. I. 1968, *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. The Journal of Finance, 23, 589–609
- Amarnath, K.N., 2004., *Statistical Methods in Consumer Credit Scoring*, Cranes Software International Ltd. Product Analyst
- Ambusaidi, M. A., He, X., Tan, Z., Nanda, P., Lu, L. F., dan Nagar, U. T., 2014., *A novel feature selection approach for intrusion detection data classification*. Trust, Security and Privacy in Computing and Communications (TrustCom), 2014 IEEE 13th International Conference on pp. 82–89.
- Baesens B, Van Gestel T, Stepanova M, Van den Poel D, Vanthienen J. *Neural network survival analysis for personal loan data*. J Oper Res Soc. 2005; 56(9): 1089 – 1098.
- Bhuyan, M. H., Bhattacharyya, D. K., dan Kalita, J. K. (2015): *Towards Generating Real-life Datasets for Network Intrusion Detection*. *IJ Network Security*, 17(6), 683–701.
- Bhaduri A., 1997., *On the formation of usurious interest rates in backward agriculture*. Camb J Econ. 1977; 1., pp. 341 – 352.
- Bramantyo Djohanputro dan Ronny Kountur, 2007, Non Performing Loan (NPL) Bank Perkreditan Rakyat (BPR), www.profi.or.id
- Chandra Dewi, 2009, *Faktor – Faktor Yang Mempengaruhi Strategi Pemberian Kredit Dan Dampaknya Terhadap Non Performing Loan (Studi Kasus pada Bank Perkreditan Rakyat di Propinsi Jawa Tengah)*, Tesis Program Pasca Sarjana Universitas Diponegoro, Semarang (tidak dipublikasikan)
- Chang, S., S. D.-o. Kim, & G. Kondo., 2015., *Predicting Default Risk of Lending Club Loans*, Machine Learning pp. 1–5.
- Chen, Y., Zeng, Y., Luo, F., dan Yuan, Z. 2016., *A New Algorithm to Optimize Maximal Information Coefficient*. PloS one, 11(6), e0157567.
- Claudia C.C., 2014, *Algoritma Klasifikasi Naïve Bayes Untuk Menilai Kelayakan Kredit*, Teknik Informatika Fakultas Ilmu Komputer, Universitas Dian

- Nuswantoro, 11 Juli.
- Dendawijaya Lukman, 2003, *Manajemen Perbankan*, Ghalia Indonesia. Jakarta.
- Dendy Saputo, 2011, *Model Credit Scoring untuk Proses Analisa Kelayakan Fasilitas Kredit Motor Menggunakan Metode Classification And Regression Tree (CART) (Studi Kasus PT. X Finance Cabang Masuk Desember 2010)*, Skripsi Fakultas Sains dan Teknologi Universitas Islam Negeri Syarif Hidayatullah, Jakarta
- Diego Alonso Gastelum Chavira, Juan Carlos Leyva Lopez, Jesus Jaime Solano Noriega, Omar Ahumada Valenzuela, Pavel Anselmo Alvarez Carrillo, 2017, A credit ranking model for a parafinancial company based on the ELECTRE-III method and a multiobjective evolutionary algorithm *Applied Soft Computing* 60 (2017) 190–201
- Eliana Angelini, Giacomo di Tollo, Andrea Roli, 2008., *A neural network approach for credit risk evaluation* *The Quarterly Review of Economics and Finance* 48, pp. 733–755
- Frendy dan Isti Sulanjari, 2014., *Pembentukan Model Credit Scoring Dengan Menggunakan Metode Bayesian Network: Stusi Kasus Permohonan Aplikasi Kredit Pemilikan Rumah (KPR)*, Fakultas Teknik Industri, Universitas Indonesia, Jakarta.
- Hand DJ, Henley WE. 1997., *Statistical Classification Methods in Consumer Credit Scoring: a Review*. *J Roy Stat Soc A Sta.* 1997; 160: 523 – 541.
- Heiat. A. 2011, *Modelling Consumer Credit Scoring Through Bayes Network*, *World Journal of Social Sciences*, Vol. 1, 132 – 141.
- Henny Leidiyana., 2013., *Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor*, *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic* 1(1)., hal. 65-76
- Huang, C.-L. , Chen, M.-C. , & Wang, C.-J. 2007., *Credit scoring with a data mining approach based on support vector machines*. *Expert Systems with Applications*, 33 , pp. 847–856 .
- Jarrow RA., 2011., *Credit market equilibrium theory and evidence: Revisiting the structural versus reduced form credit risk model debate*. *Financ Res Lett.*, 8 (1)., pp. 2 – 7.
- Jonathan N. Crook, David B. Edelman, Lyn C. 2007., *Recent developments in consumer credit risk assessment* *Thomas*, *European Journal of Operational Research* 183., pp. 1447–1465.
- Lean Yu, Shouyang Wang, Kin Keung Lai, 2008., *Credit risk assessment with a multistage neural network ensemble learning approach*, *Expert Systems with Applications* 34, pp. 1434–1444
- Lee, T. S., Chiu, C. C., Lu, C. J., & Chen, I. F. 2002., *Credit scoring using the hybrid neural discriminant technique*. *Expert Systems with Application*, 23(3), pp. 245–254.
- Li, S.-T., Shiue, W., & Huang, M.-H., 2006., *The Evaluation of consumer loans using support vector machines*. *Expert Systems with Applications*, 30, pp. 772–782.
- Kasmir, 2001, *Bank Dan Lembaga Keuangan*, Edisi Revisi. Penerbitan: PT. Raja Cerfindo Persada, Jakarta
- M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, 2016., *Building an intrusion detection system using a filter-based feature selection algorithm*, *IEEE*

- transactions on computers, vol. 65, no. 10, pp. 2986–2998.
- M. Doumpos, K. Kosmidou, G. Baourakis, C. Zopounidis, 2002., *Credit risk assessment using a multicriteria hierarchical discrimination approach: A comparative analysis*, European Journal of Operational Research 138, pp. 392–412
- Matteo Accornero, Giuseppe Cascarino, Roberto Felici, Fabio Parlapiano, Alberto Maria Sorrentino, 2017., *Credit risk in banks ' exposures to non-financial firms*, Eur Financ Manag, pp. 1 – 17.
- Min, J. H. , & Lee, Y.-C., 2005., *Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters*. Expert Systems with Applications, 28, pp. 603–614.
- Neto Rosalvo, Paulo Jorge Adeodato, dan Ana Carolina Salgado, 2016, *A Framework for Data Transformation in Credit Behavioral Scoring Applications Based on Model Driven Development*, Journal of Expert Systems with Applications, November.
- Ni Wayan Vany Ekaulandari, A.A.N.B Dwirandra, 2013, *Pengaruh Penaksiran Resiko, Informasi Dan Komunikasi, Aktivitas Pengendalian, Pemantauan, Lingkungan Pengendalian Pada Efektivitas Sistem Pemberian Kredit*, E-Jurnal Akuntansi Universitas Udayana 4.3., hal. 585-604
- Nie, G., Rowe, W., Zhang, L., Tian, Y., & Shi, Y. 2011., *Credit card churn forecasting by logistic regression and decision tree*. Expert Systems with Applications, 38, pp. 15273–15285 .
- Novita Ikasari, 2014, *Credit Decision Support Methodology for Micro, Small and Medium Enterprises (MSMEs): Indonesian Cases*, thesis for the Degree of Doctor of Philosophy of Curtin University, June
- Noura Metawa, M. Kabir Hassan, Mohamed Elhoseny, 2017., *Genetic algorithm based model for optimizing bank lending decisions*, Expert Systems With Applications 80., pp. 75–82
- Othman, Z. A., Bakar, A. A., dan Etubal, I., 2010., *Improving signature detection classification model using features selection based on customized features*. Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on IEEE., pp. 1026–1031).
- Putra, S. D., Ahmad, A. S., dan Sutikno, S. 2016., *DPA-countermeasure with knowledge growing system*. Electronics and Smart Devices (ISESD), International Symposium on IEEE pp. 16–20
- Pemerintah Republik Indonesia. 1998. Undang-Undang Republik Indonesia Nomor 10 Tentang Perbankan, Jakarta, Indonesia: Government of Indonesia.
- Qi Zhang, Jue Wang, Aiguo Lu, Shouyang Wang, Jian Ma, *An Improved SMO Algorithm for Financial Credit Risk Assessment—Evidence from China's banking*, Neurocomputing 2017, doi: 10.1016/j.neucom.2017.07.002
- Qinwei Chi, Wenjing Li, 2017., *Economic policy uncertainty, credit risks and banks' lending decisions: Evidence from Chinese commercial banks*, China Journal of Accounting Research 10., 33–50
- Rivai, Veithzal.,&Veithzal, Andria Permata. 2006. *Credit Management Handbook*. Jakarta : Raja GrafindoPersada.
- Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., Lander, E. S., Mitzenmacher M., Sabeti, P., C. 2011.,

- Detecting novel associations in large data sets. science*, 334(6062), 1518–1524.
- Robynson Amseke, Edi Winarko, 2014, *Aplikasi Algoritma CBA untuk Klasifikasi Resiko Pemberian Kredit (Studi kasus: PT. Telkom CDC Sub Area Kupang)* IJCCS, Vol.8, No.2, July 2014, pp. 121~132
- Robert R. Trippi dan Efraim Turban, 1992, *Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real World Performance*, McGraw-Hill, Inc. New York, NY, USA
- Sumari, A.D.W, 2010., *Sistem Berpengetahuan-Tumbuh: Satu Perspektif Baru Dalam Kecerdasan Tiruan*, Disertasi Doktor Teknik Elektro dan Informatika, Institut Teknologi Bandung, 2010.
- Tânia S. H. Gonçalves, Fernando A. F. Ferreira, Marjan S. Jalali & Ieva, Meidutė-Kavaliauskienė 2015., *An idiosyncratic decision support system for credit risk analysis of small and medium-sized enterprises*, Technological and Economic Development of Economy, 02 Nov 2015.
- Teguh Wahyono, Ariya Dwika Cahyono, 2015, *Pengembangan Model Mitigasi Resiko Kredit Berbasis Komputasional Untuk Meningkatkan Kemamouan Manajemen Resiko Bagi Koperasi*, Jurnal Sistem Komputer– Vol. 5, No 1, Mei 2015.
- Tom Cronje dan Apriani Atahau, 2017, *Bank Lending – Theory and Practice 2e*, McGraw- Hill Education.
- Twala, B. 2010., *Multiple classifier application to credit risk assessment*. Expert Systems with Applications, 37, 3326–3336.
- Yufei Xia, Chuanzhe Liu, YuYing Li, Nana Liu, 2017, *A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring* , Expert Systems With Applications 78, 225–241
- Zhang, Y. , Jia, H. , Diao, Y. , Hai, M. , & Li, H., 2016, *Research on Credit Scoring by Fusing Social Media Information in Online Peer-to-Peer Lending*. Procedia Computer Science, 91 , 168–174 .
- Zhu, Y., Liang, J., Chen, J., dan Ming, Z. (2017): *An improved NSGA-III algorithm for feature selection used in intrusion detection*. Knowledge-Based Systems, 116, 74–85.