

## model penilaian risiko kredit bawaan jangka pendek pohon berbasis membangun keputusan

Yung-Chia Chang, Kuei-Hu Chang, Heng-Hsuan Chu & Lee-Ing Tong

**Untuk mengutip artikel ini:** Yung-Chia Chang, Kuei-Hu Chang, Heng-Hsuan Chu & Lee-Ing Tong (2016) Menetapkan keputusan jangka pendek model penilaian risiko kredit bawaan berbasis pohon, Komunikasi Statistik - Teori dan Metode, 45:23, 6803- 6815, DOI: [10.1080 / 03610926.2014.968730](https://doi.org/10.1080/03610926.2014.968730)

**Untuk link ke artikel ini:** <https://doi.org/10.1080/03610926.2014.968730>



versi penulis diterima diposting online: 25 Jan 2016.

Diterbitkan online: 3 September 2016.



Mengirimkan artikel Anda ke jurnal ini



pandangan Pasal: 329



Lihat artikel terkait



Data View Crossmark



Mengutip artikel: 3 View mengutip artikel





## model penilaian risiko kredit pendek termdefault pohon berbasis membangun keputusan

Yung-Chia Chang <sup>Sebuah</sup>, Kuei-Hu Chang <sup>•</sup>, Heng-Hsuan Chu <sup>Sebuah</sup>, dan Lee-Ing Tong <sup>Sebuah</sup>

<sup>Sebuah</sup> Jurusan Teknik Industri dan Manajemen, Universitas Nasional Chiao Tung, Hsinchu, Taiwan;

<sup>•</sup> Departemen Ilmu Manajemen, ROC Akademi Militer, Kaohsiung, Taiwan

### ABSTRAK

model penilaian risiko kredit tradisional tidak mempertimbangkan waktu tor tor; mereka hanya memikirkan apakah pelanggan akan default, tapi tidak thewthen ke default. Hasilnya tidak dapat memberikan manajer untuk membuat keputusan maksimum profit. Sebenarnya, bahkan jika default pelanggan, lembaga resmi dalam hal keuangan masih bisa memperoleh keuntungan dalam beberapa kondisi. Saat ini, sebagian besar penelitian menerapkan Cox proportional hazards model ke model credit scoring mereka, memprediksi waktu ketika seorang pelanggan yang paling mungkin untuk default, untuk memecahkan masalah penilaian risiko kredit. Namun, dalam rangka untuk sepenuhnya memanfaatkan kemampuan dinamis sepenuhnya dari Cox Model ARDS proporsional haz-, variabel makroekonomi waktu bervariasi yang diperlukan yang melibatkan pengumpulan data yang lebih maju. Sejak jangka pendek kasus bawaan adalah orang-orang yang membawa kerugian besar bagi lembaga keuangan, bukannya memprediksi saat kredit akan default, manajer pinjaman lebih antar ested dalam mengidentifikasi aplikasi yang whichmay standar dalam waktu singkat ketika menyetujui aplikasi pinjaman. Makalah ini mengusulkan sebuah model penilaian risiko kredit bawaan jangka pendek berbasis pohon keputusan untuk menilai risiko kredit. Tujuannya adalah untuk menggunakan pohon keputusan untuk menyaring default jangka pendek untuk menghasilkan model yang sangat akurat yang bisa dis pinjaman standar tinguish. Makalah ini mengintegrasikan diagregasi bootstrap (Bagging) dengan minoritas sintesis teknik over-sampling (memukul) ke dalam model risiko kredit untuk meningkatkan stabilitas pohon keputusan dan kinerjanya pada data yang tidak seimbang. Akhirnya, kasus nyata dari data pinjaman kecil dan mediumenterprise yang telah ditarik dari lembaga keuangan lokal yang terletak di Taiwan disajikan untuk lebih menggambarkan pro yang ditimbulkan pendekatan. Setelah membandingkan hasil yang diperoleh dari pendekatan yang diusulkan dengan regresi logistik dan Cox model proporsional bahaya, ditemukan bahwa tingkat klasifikasi recall dan tingkat sion preci- dari model yang diusulkan jelas lebih unggul dari regresi logistik dan Cox proportional hazards model .

### PASAL SEJARAH

Menerima •• Januari •••• Diterima

•• September ••••

### KATA KUNCI

Bootstrap menggabungkan;  
model penilaian risiko kredit;  
pohon keputusan; bawaan jangka pendek.

### KLASIFIKASI SUBJECT

#### MATEMATIKA

•• F••; ••B••

## 1. Pendahuluan

Pinjaman adalah salah satu perilaku komersial perdagangan umum dalam masyarakat modern, dan receive- ing bunga adalah salah satu cara utama bahwa lembaga keuangan mendapatkan keuntungan. Taiwan bank lokal pendapatan bunga bersih adalah 63% dari total laba bersih di tahun 2011 (Bank Sentral Republik China, 2011 ). Melalui penilaian risiko kredit, lembaga keuangan dapat memperkirakan keuntungan atau kerugian potensial, biasanya mengandalkan penilaian onhuman di masa lalu. Saat ini, ada





banyak lembaga keuangan memprediksi probabilitas default melalui model penilaian risiko kredit. Dengan data historis dan teknik statistik, model penilaian risiko kredit berusaha untuk membedakan karakteristik yang berbeda dari peminjam untuk kenakalan dan default (Mester, 1997). Saat ini, ada banyak penelitian difokuskan pada model penilaian risiko kredit, tetapi sebagian besar model penilaian didasarkan pada struktur peramalan apakah peminjam akan default atau tidak, tanpa memperhatikan titik waktu dari default (Altman, 1968; Thomas, 2000). pembayaran pinjaman adalah proses kemajuan dari waktu ke waktu. Mari kita mendefinisikan pendek termdefault sebagai default terjadi dalam jangka waktu tertentu (misalnya, standar dalam 1 tahun untuk pinjaman 2 tahun) setelah pinjaman disetujui. Setiap bawaan yang terjadi lebih dari waktu yang disebut standar jangka panjang. Untuk lembaga keuangan, pendek-termdefault biasanya menyebabkan lebih banyak kerugian daripada satu jangka panjang dengan ukuran yang sama, sehingga harus alamat dengan risiko yang lebih tinggi. Hal ini karena jangka waktu yang lama pinjaman dilunasi, lebih modal dan bunga lembaga keuangan dapat pulih. Oleh karena itu, mungkin terjadi bahwa bahkan ketika pinjaman adalah default, lembaga keuangan masih dapat keuntungan fromproviding pinjaman itu. Tujuan dari lembaga keuangan adalah laba tomaximize; dengan demikian, analisis risiko juga harus take profit menjadi pertimbangan. 2004).

Untuk melengkapi kekurangan dari model biner risiko kredit penilaian yang dikembangkan dalam beberapa tahun terakhir, beberapa sarjana telah mulai menganalisis waktu pembayaran pinjaman. Narain (1992) Adalah salah satu peneliti awal yang memperkenalkan analisis survival ke dalam risiko kredit assess- Model ment. Sebagian besar analisis survival untuk model penilaian risiko kredit mengadopsi Cox metode proportional hazards (Cox, 1972) Untuk menetapkan waktu model penilaian risiko kredit (Tong et al., 2012). Secara tradisional, penilaian risiko menggunakan Cox metode proportional hazards mengalami beberapa kesulitan dan gagal untuk meningkatkan akurasi prediksi. Namun, ada masih belum menyimpulkan hasil mengenai dampak peristiwa default. Misalnya, dalam kredit korporasi, Shumway (2001) Menemukan bahwa lamanya waktu suatu perusahaan didirikan tidak secara statistik signifikan dalam model hazard mereka. Namun, Eklund et al. (2001) Menunjukkan bahwa sebuah perusahaan dengan waktu yang lebih singkat didirikan memiliki probabilitas yang lebih tinggi dari standar perusahaan. Di Bertawanan con, Bonfim (2009) Menunjukkan temuan yang berbeda lengkap dibandingkan dengan apa yang ditemukan oleh Eklund et al. (2001), Yaitu, suatu perusahaan yang didirikan lagi memiliki probabilitas yang lebih tinggi dari default perusahaan. Di sisi lain, bahaya proporsional Cox model yang mengasumsikan bahwa semua peristiwa pada akhirnya akan terjadi; yaitu, semua pemberi pinjaman pada akhirnya akan default, tapi itu tidak benar dalam praktek (Tong et al., 2012). Banasik et al. (1999) Menggunakan Cox proportional hazard model yang membangun kelangsungan hidup

model untuk pinjaman pribadi; themodel memiliki akurasi tinggi pada prediksi default untuk kasus dalam waktu 1 tahun, tapi tampil buruk untuk kasus-kasus lebih dari 1 tahun. Baesens et al. (2005) Digunakan jaringan saraf untuk membangun model credit scoring dan membuktikan bahwa kinerja model sedikit lebih baik dari Cox model yang bahaya proporsional. Duffie et al. (2007) Menggunakan metode time series untuk membangun sebuah model penilaian risiko sementara mengintegrasikan kovariat makroekonomi untuk model untuk memprediksi probabilitas default untuk perusahaan diperdagangkan publik di Amerika Serikat dalam jangka waktu tertentu; Model ini juga memiliki akurasi tinggi pada prediksi untuk standar terjadi dalam waktu 1 tahun.

Keuntungan themain pohon keputusan adalah bahwa hal itu tidak memerlukan variabel yang harus mematuhi asumsi statistik yang ketat. Alih-alih memprediksi lamanya waktu kelangsungan hidup untuk setiap kasus, seperti apa bahaya proporsional Cox model yang tidak, pohon keputusan dapat diterapkan untuk menyaring kasus pendek termdefault, yang harus addressedwith risiko yang lebih tinggi. Pendekatan ini mengatasi kelemahan bahwa model risiko kredit biner tradisional tidak mempertimbangkan



waktu default. Akurasi prediksi juga bisa lebih baik daripada hanya menggunakan survival analysis. Namun, pohon keputusan memiliki dua kekurangan utama. Pertama, pohon keputusan bukan classifier stabil: struktur cabang pohon keputusan sensitif terhadap pelatihan sampel; sampel yang berbeda, bahkan jika sampel dari populasi yang sama, biasanya menghasilkan struktur cabang yang berbeda. Oleh karena itu, keakuratan pohon keputusan sangat dipengaruhi oleh pemilihan sampel, gagal untuk menjamin kemampuan diskriminasi stabil pohon keputusan. Untuk secara efektif mengatasi masalah ini, Breiman (1996) mengusulkan *diaggresi bootstrap* (menggunakan akronim *Bagging*), melalui perluasan ukuran sampel yang ada dan meningkatkan jumlah pohon keputusan.

Kedua, seperti kebanyakan pengklasifikasi, pohon keputusan tidak dapat secara efektif menangani data tidak seimbang, di mana jumlah mata pelajaran dalam satu kategori (disebut kelas mayoritas) secara signifikan lebih dari yang lain. Pohon keputusan cenderung untuk membedakan semua mata pelajaran untuk kelas mayoritas untuk data tidak seimbang. Data dari aplikasi pinjaman adalah contoh khas dari data yang tidak seimbang karena kasus standar biasanya hanya menyumbang sebagian kecil dari keseluruhan data. Dalam rangka meningkatkan kemampuan pohon keputusan untuk menangani data tersebut, salah satu cara adalah dengan menerapkan sintesis Minoritas teknik *over-sampling* (memukul) yang dikembangkan oleh Chawla et al. (2002) Untuk menyeimbangkan jumlah mata pelajaran di antara berbagai kategori dalam sampel pelatihan. Berbeda dengan pendekatan hanya duplikasi data yang ada, memukul membesar dataset dengan menciptakan data baru dari yang sudah ada untuk mengaburkan batas kategori yang berbeda dari sampel. Hal ini dapat mencegah pengklasifikasi dari *overfitting* data dan juga dapat meningkatkan daya diskriminasi dari classifier.

Dalam rangka untuk secara efektif memecahkan masalah tersebut di atas, makalah ini mengusulkan sebuah pendekatan dengan mengintegrasikan *Bagging* dan memukul pohon keputusan untuk membangun model penilaian risiko kredit bawaan jangka pendek. Tujuan dari model penilaian kredit ini adalah untuk mengidentifikasi kasus bawaan jangka pendek dari aplikasi semua peminjam. Model ini didasarkan pada pohon keputusan C4.5. Metode *Bagging* diterapkan untuk meningkatkan kemampuan membedakan dari pohon keputusan. Memukul diterapkan untuk memperbesar data dari kelas minoritas (yaitu, kasus default) dalam sampel pelatihan untuk menyeimbangkan ukuran default dan data non standar saat pelatihan classifier.

Sisa kertas diatur sebagai berikut: Sec. 2 ulasan penelitian terkait tentang kertas-C4.5 ini keputusan pohon dan *diaggresi bootstrap*, memukul, dan Cox model yang bahaya proporsional. Jangka pendek risiko kredit bawaan model penilaian berbasis pohon keputusan yang mengintegrasikan *Bagging* dan memukul diusulkan di Sec. 3. Kasus aktual dari data pinjaman andmedium usaha kecil, yang disediakan oleh lembaga keuangan yang berlokasi di Taiwan, diadopsi, dan beberapa perbandingan dengan pendekatan yang terdaftar dibahas di Sec. 4. Bagian akhir membuat kesimpulan.

## 2. Relatedworks

### 2.1. pohon keputusan C4.5 dan *diaggresi bootstrap*

Quinlan (1993) Mengembangkan C4.5 pohon keputusan classifier untuk menentukan strategi percabangan oleh jumlah pengurangan entropi. Pohon keputusan C4.5 merupakan versi perbaikan dari Quinlan ini (1986) Pohon keputusan ID3. Keuntungan terbesar adalah dengan menggunakan rasio keuntungan untuk menggantikan keuntungan *Information* dalam memilih variabel bercabang untuk menghindari *overfitting*.

Pohon keputusan C4.5 digunakan rasio keuntungan sebagai kriteria bercabang; Langkah pertama algoritma adalah untuk menghitung rasio keuntungan dari masing-masing variabel dan kemudian cabang di satu dengan rasio keuntungan tertinggi. Data tersebut kemudian dibagi menjadi beberapa subkelompok non tumpang tindih sesuai dengan kondisi split (nilai variabel bercabang). Setiap subkelompok kemudian bercabang



demikian. Rasio keuntungan dihitung dengan Persamaan (1) (Quinlan, 1993):

$$\text{Rasio Gain} = \frac{\text{gain Informasi} (Y/X)}{\text{Informasi split} (Y/X)} = \frac{I(Y) - I(Y/X)}{I(X)} \quad (1)$$

dimana  $X$  adalah variabel dan  $Y$  adalah atribut  $X$ .

gain informasi digunakan untuk mengukur jumlah informasi diperoleh dengan percabangan pada variabel. Informasi perpecahan standarisasi tingkat dispersi dalam variabel. Untuk variabel acak  $P$  dengan  $m$  hasil, jumlah informasi (entropi), dinotasikan dengan  $AKU(P)$ , dihitung sebagai berikut:

$$I(P) = - \sum_{j=1}^m p_j \log_2 p_j \quad (2)$$

dimana  $p_j$  adalah probabilitas hasil  $j$ , untuk  $j = 1, \dots, m$ .

Entropi bersyarat  $Y$  diberikan  $X$  didefinisikan sebagai:

$$I(Y/X) = \sum_j P(X = v_j) \cdot I(Y/X = v_j) \quad (3)$$

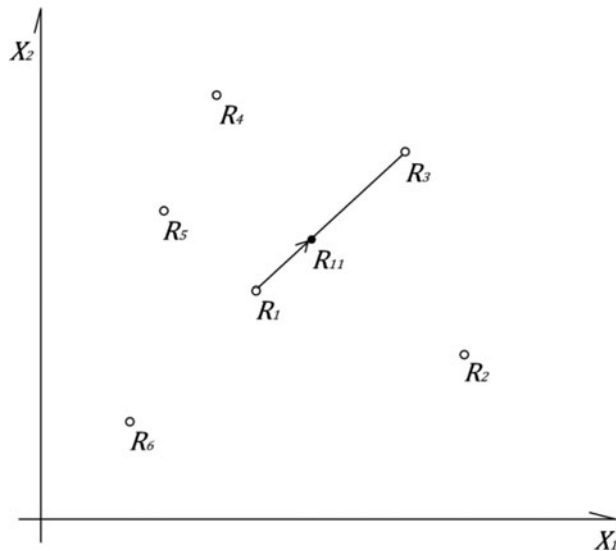
dimana  $I(Y/X = v_j)$  adalah informasi posterior  $Y$  setelah mengamati nilai-nilai  $X = v_j$ .

Breiman (1996) Menunjukkan bahwa proses klasifikasi pohon keputusan tidak stabil; perubahan kecil pada hasil sampel dalam perubahan signifikan pada pohon keputusan. Untuk mengatasi masalah ini, Breiman (1996) Karena itu diusulkan Bagging untuk meningkatkan efektivitas klasifikasi pohon keputusan. Konsep Bagging adalah untuk membangun beberapa pohon keputusan untuk menghindari dampak dari tidak stabil klasifikasi pohon keputusan tunggal. Sampel asli pertama kali digandakan melalui bootstrap method into multiple sampel baru yang dapat mewakili sampel asli dan kemudian melatih pohon keputusan untuk setiap sampel baru. Sampel baru yang dihasilkan oleh metode bootstrap sedikit berbeda antara satu sama lain; sehingga pohon-pohon keputusan dilatih tidak akan sama. Hasil klasifikasi akhir ditentukan oleh keputusan mayoritas pohon keputusan yang dihasilkan. Breiman (1996) Membuktikan bahwa metode Bagging secara signifikan dapat meningkatkan kemampuan klasifikasi pohon keputusan.

## 2.2. memukul

Chawla et al. (2002) Mengembangkan memukul, teknologi data sintetik untuk memecahkan lem masalah. Safe\_mode data tidak seimbang. Teknologi ini tidak over-sampel kelas minoritas dengan ment replace- dengan menciptakan "sintetik" data sesuai dengan karakteristik kelas minoritas untuk menghindari overfitting ketika melatih classifier a.

Aturan memukul adalah untuk pertama secara acak memilih  $k$  tetangga terdekat dari setiap kasus minoritas kelas (seperti kasus default)  $R_{sayo}$  berdasarkan jarak Euclidean. Acak memilih sebuah kasus, yang disebut  $R_i$  antara tetangga tersebut  $R_{sayo}$ . kasus  $R_{sayo}$  dan  $R_i$  bisa membentuk vektor  $R_{sayo} R_i$ . Menghasilkan nomor acak  $x$  antara 0 dan 1. kasus default yang baru dihasilkan  $R_k$  adalah sama dengan  $R_{sayo}$  ditambah vektor  $x \cdot R_{sayo} R_i$  (Gambar 1). Ulangi langkah-langkah tersebut sampai kondisi keseimbangan terpenuhi. Menghadapi data yang berisi variabel diskrit, perhitungan dikan dis-Euclidean antara dua kasus bawaan perlu disesuaikan. Pertama, menghitung deviasi standar dari setiap variabel kontinu di  $R_{sayo}$ ; biarkan median ini standar deviasi akan Med. Saat menghitung jarak Euclidean dari  $R_{sayo}$  dan  $R_i$  untuk variabel diskrit yang nilainya dalam dua kasus tidak sama, menggunakan Med sebagai jarak antara dua titik; sebaliknya, mengatur jarak ke nol. Ketika menghasilkan kasus baru, menggunakan nilai yang muncul paling di antara  $k$  tetangga sebagai nilai untuk variabel diskrit.



**Gambar •** Diagram skematik dari memukul.

### 2.3. Cox hazardsmodel proporsional

Model kelangsungan hidup, dibandingkan dengan analisis regresi, menganggap pengaruh peristiwa sampel yang belum terjadi pada themodel parameter (seperti peminjam membayar tepat waktu tanpa default ketika kontrak berakhir). Model survival mengasumsikan suatu peristiwa akan terjadi di masa depan. Saat ini, sebagian besar penelitian menggunakan Cox proportional hazard model yang membangun model survival (Cox, 1972). Cox proportional hazards model yang mengasumsikan bahwa fungsi hazard

1 pada waktu  $t$  dihitung sebagai Persamaan (4):

$$\lambda(t/X) = \lambda_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p) = \lambda_0(t) \exp(\beta' X) \quad (4)$$

dimana  $\lambda_0(t)$  adalah fungsi dasar hazard,  $X$  adalah variabel yang mempengaruhi lamanya waktu hidup sampel, dan  $\beta$  adalah koefisien yang sesuai dengan masing-masing variabel.

Nilai dari  $\beta$  dapat diperoleh dengan metode kemungkinan parsial. Menganggap  $Y_{saya}$  lamanya waktu yang diamati sampel *saya*; yaitu, waktu berlalu dari awal sampel *saya* sampai baik terjadi peristiwa atau akhir periode observasi adalah  $Y_{saya}$ . Selain itu, biarkan  $C_{saya}$  mengungkapkan apakah suatu peristiwa telah terjadi sampel *saya* selama periode pengamatan; jika telah terjadi peristiwa maka

$C_i = 1$ ; jika tidak,  $C_i = 0$ . Fungsi kemungkinan parsial dinyatakan sebagai:

$$L(\beta) = \prod_{i: C_i = 1} \frac{\theta_{saya}}{\sum_{j: Y_j > Y_i} \theta_j} \quad (5)$$

dimana  $\theta_j = \beta' X_j$ .

Keuntungan dari Cox Model bahaya proporsional adalah bahwa hal itu tidak memerlukan satu untuk mengasumsikan distribusi statistik saat acara terjadinya.

### 3. pendek termdefault model penilaian risiko kredit berbasis pohon keputusan Usulan

Langkah pertama untuk membangun model kami adalah untuk mengumpulkan dan menyusun informasi. Kita harus menggunakan metode sonable Reason untuk memilih variabel yang berpengaruh (disebut seleksi fitur) untuk memprediksi probabilitas default jangka pendek.

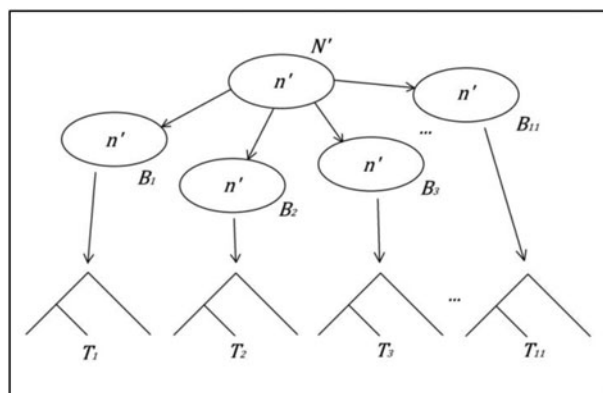
Makalah ini digunakan regresi logistik sebagai pilihan fitur

metode; variabel respon adalah terjadinya peristiwa bawaan jangka pendek, variabel secara bebas adalah orang-orang yang signifikan pengaturan  $\alpha = 0.01$ . Selain itu, kita perlu untuk menghasilkan sampel pelatihan untuk pelatihan model yang. Sejak kasus bawaan jangka pendek beberapa dibandingkan dengan kasus non default, kebanyakan pengklasifikasi tidak dapat secara efektif mengidentifikasi kasus ini langsung karena masalah ketidakseimbangan kelas. Classifier yang gagal untuk mengidentifikasi default tidak akan memiliki aplikasi kal prakti-. Oleh karena itu, kertas yang digunakan memukul ini untuk selama-sampel kasus bawaan jangka pendek untuk menyeimbangkan data dua kelas; untuk lebih tepatnya, oleh memukul, jumlah total kasus bawaan jangka pendek diperbesar sehingga rasio jumlah kasus standar jangka pendek dengan yang non default adalah 1:  $\theta$ , dimana  $\theta$  adalah nomor alam ditentukan oleh pembuat keputusan. Misalkan sampel yang  $N$  memiliki  $n$  kasus pinjaman, yang terdiri dari  $r$  kasus bawaan jangka pendek dan  $(n - r)$  kasus non default. Membiarkan  $\theta$  menjadi jumlah kasus default jangka pendek setelah menerapkan memukul. Tujuan kami untuk menyeimbangkan data dalam dua kelas adalah untuk membuat  $\theta \sim (n - r)$ . Dengan cara itu, semua kasus default non tetap dalam sampel pelatihan, sehingga semua informasi dari kasus non standar masih dipertahankan dalam sampel pelatihan bukan kehilangan beberapa atas bawah-sampling. Perhatikan bahwa  $\theta$  disesuaikan per data yang dikumpulkan. Misalnya, satu dapat menyesuaikan  $\theta$  1, 2, 3, atau 4 untuk membuat jumlah kasus standar jangka pendek untuk yang non standar sebagai 1: 1, 1: 2, 1: 3, atau 1: 4. Hal ini dimungkinkan untuk menggunakan metode trial-and-error untuk memilih strategi keseimbangan terbaik untuk dataset yang diberikan. Karena pengalaman kami dalam bekerja dengan data pinjaman kredit, pengaturan  $\theta$  sebagai 2, yaitu, membuat jumlah kasus standar jangka pendek untuk jumlah kasus non standar sebagai 1: 2, sering dapat memberikan hasil yang baik.

Setelah menyelesaikan pengaturan data dan balancing data, kita dapat mulai untuk melatih fier classi- untuk penggunaan cases. We pendek termdefault  $N$  untuk menunjukkan sampel setelah menyesuaikan jumlah kasus pendek termdefault oleh memukul dan penggunaan  $n$  untuk mewakili jumlah kasus di  $N$ . Pertama, menggunakan metode bootstrap untuk menghasilkan 11 sampel bootstrap,  $B_1, B_2, \dots, B_{11}$ , masing-masing memiliki  $n$ .

kasus. Untuk setiap sampel bootstrap  $B_{\text{saye}}$ , menggunakan algoritma C4.5 untuk melatih pohon keputusan  $T_{\text{saye}}$  untuk  $i = 1, \dots, 11$ . Gambar 2 menggambarkan konsep dari proses ini.

Jangka pendek risiko kredit bawaan model penilaian didirikan adalah 11 pohon keputusan dilatih ( $T_1, T_2, \dots, T_{11}$ ). Ketika sedang diterapkan untuk menentukan kapan kasus baru akan cenderung ke default di waktu singkat, hanya plug-in data kasus yang ke masing-masing pohon keputusan secara independen. Karena setiap pohon keputusan akan menghasilkan satu hasil, 11 hasil diklasifikasikan akan diperoleh. Hasil akhir dari kasus baru ini ditentukan oleh hasil mayoritas pohon keputusan tersebut. Artinya, jika lebih dari setengah dari pohon keputusan menentukan kasus seperti jangka pendek default, maka hal ini diidentifikasi sebagai kasus berisiko tinggi; jika tidak, itu adalah bawaan non



**Gambar 2.** Ilustrasi agregasi bootstrap dan pohon keputusan C<sub>4.5</sub>.

**Tabel 4.** Kebingungan matriks.

	diprediksi negatif	diprediksi positif
aktual negatif	TN	FP
sebenarnya positif	FN	TP

huruf dengan risiko yang relatif rendah. Misalnya, jika pohon keputusan ( $T_1, T_2, \dots, T_{11}$ ) memiliki berbeda- hasil ent membedakan (non default, standar jangka pendek, non default, non default, non default, non default, standar jangka pendek, standar jangka pendek, non default, non default, non default); ada delapan kasus bawaan non dalam hasil diskriminasi dan tiga kasus default. Sejak delapan pohon keputusan mengkategorikan kasus ini sebagai salah satu non default, model ini menyimpulkan itu adalah kasus non default. Breiman ( 1996 ) Menunjukkan bahwa ketika jumlah bootstrap sam- prinsip keuangan jauh lebih dari 10, biaya operasional tambahan dari pohon keputusan akan lebih besar daripada manfaat marjinal bootstrap. Selanjutnya, untuk menghindari situasi imbang (hasil memiliki lima non default dan lima jangka pendek default), makalah ini digunakan 11 sampel bootstrap.

ukuran kinerja yang digunakan dalam makalah ini adalah mengingat dan presisi. Biarkan TP menjadi jumlah positif (jangka pendek default, dalam makalah ini) kasus diidentifikasi dengan benar, FN menjadi jumlah kasus positif yang salah diklasifikasikan sebagai negatif (non-default, dalam hal ini), FP menjadi jumlah kasus negatif tidak benar diklasifikasikan sebagai positif, dan TN akan jumlah kasus negatif dengan benar diklasifikasikan sebagai negatif. confusionmatrix yang ditunjukkan pada Tabel 1 (Chawla, et al.,

2002 ). Perhitungan recall dan presisi didasarkan pada Persamaan ( 6 ) Dan ( 7 ).

$$\text{ingat} = TP \frac{\text{---}}{TP + FN} \quad (6)$$

$$\text{presisi} = TP \frac{\text{---}}{TP + FP} \quad (7)$$

Model yang diusulkan diatur dalam prosedur berikut:

Langkah 1. Mengumpulkan dan mengatur data. Gunakan regresi logistik untuk memilih indepen- signifikan penyok variabel yang memiliki pengaruh yang lebih besar untuk membuat kasus seperti jangka pendek default, di  $\alpha = 0.01$ .

Langkah 2. Gunakan memukul untuk menyeimbangkan jumlah default dan standar kasus jangka pendek di pelatihan sampel. sampel pelatihan mengakibatkan disebut sampel seimbang. Biarkan jumlah kasus standar jangka pendek mencapai setengah dari jumlah kasus non default dalam sampel yang seimbang.

Langkah 3. Menghasilkan 11 sampel bootstrap dengan ukuran yang sama berdasarkan sampel yang seimbang. Langkah 4. Gunakan algoritma C4.5 untuk melatih satu pohon keputusan untuk setiap sampel bootstrap. Langkah 5. Uji themodel menggunakan data pengujian. Hasil klasifikasi ditentukan oleh themajority

hasil dari 11 pohon keputusan dilatih di Langkah 4.

#### 4. verifikasi numerik dan perbandingan

Pada bagian ini, makalah ini menggunakan data pinjaman riil usaha kecil dan menengah (UKM) untuk memverifikasi kemampuan klasifikasi model yang diusulkan. Makalah ini juga membandingkan hasil dengan menggunakan regresi logistik dan Cox proportional hazard model untuk menangan data pinjaman kecil dan menengah.



**Tabel 1.** Jumlah kasus di semua sampel.

	Semua	Jumlah jangka pendek kasus bawaan	Persentase kasus bawaan jangka pendek pada semua sampel
sampel asli	22.433	433	1.9%
sampel yang seimbang	22.433	433	1.9%

#### 4.1. Sumber data

Pengujian dan pelatihan sampel dari model penilaian risiko adalah data beban pro-vided oleh lembaga keuangan yang berlokasi di Taiwan yang dikumpulkan antara November 2001 dan Desember 2010. Pinjaman ini dibuat untuk UKM. Sebagian besar pinjaman ini adalah pinjaman kredit jangka pendek. Median durasi pinjaman adalah 24 bulan. Sebanyak 12.514 kasus pinjaman dilakukan selama waktu pengumpulan data, di antaranya 932 (atau 7,4%) kasus gagal bayar dalam waktu 1 tahun setelah mereka diberikan. Kami nama kasus-kasus seperti kasus bawaan jangka pendek. Menggunakan memukul, 4660 kasus bawaan jangka pendek diciptakan yang membuat total ber NUM kasus bawaan jangka pendek sebagai 5592, akuntansi untuk 32,6% dari semua kasus (seperti yang ditunjukkan pada

#### Meja 2 ).

Setiap aplikasi pinjaman berisi hasil dari latar belakang investigasi perusahaan, data keuangan berdasarkan laporan keuangan perusahaan, dan indeks yang menilai situasi lingkungan ekonomi secara keseluruhan sebelum lembaga keuangan memberikan pinjaman. Lebih dari 30 variabel yang terkait dengan satu aplikasi pinjaman. Setelah screening variabel melalui regresi logistik untuk slimdown jumlah variabel yang dibutuhkan, 13 variabel berikut ditemukan signifikan di bawah  $\alpha = 0.01$ : **Sejarah Perusahaan (usia)**: sesuai dengan waktu yang ditetapkan dan pengembangan perusahaan, dibagi

menjadi empat tingkat, dinilai oleh penilai pinjaman profesional di lembaga keuangan. tingkat Factorymanagement ( *factory\_management* ) mengukur jika pabrik peminjam adalah baik-

dikelola, dibagi menjadi dua tingkat, dinilai oleh penilai kredit profesional. Rencana bisnis ( *rencana\_bisnis* ): menunjukkan jika peminjam memiliki rencana bisnis yang jelas dan suara,

dibagi menjadi lima tingkat, dinilai oleh penilai kredit profesional. Reputasi dari pejabat eksekutif di industri ( *Latar Belakang* ): menunjukkan reputasi

executive officer dari perusahaan dalam industri, dibagi menjadi lima tingkat, dinilai oleh penilai pinjaman sional guru besar. status kredit dari bank lain ( *bank\_evaluation* ): menunjukkan apakah peminjam memiliki buruk

kredit di bank lain pada catatan resmi, dibagi menjadi dua tingkat. reputasi perusahaan dengan pesaing ( *competitor\_evaluation* ), menunjukkan hasil peer

evaluasi, dibagi menjadi dua tingkat, dinilai oleh penilai pinjaman setelah melakukan penyelidikan untuk perusahaan lain dalam industri yang sama.

tingkat aktiva bersih pejabat eksekutif atau penjamin ( *modal* ): menunjukkan tingkat aktiva bersih

dari pejabat eksekutif atau penjamin, dibagi menjadi lima tingkat. Rasio kewajiban ( *liability\_ratio* ): Rasio kewajiban dihitung dengan membagi total kewajiban dengan jumlah

aset, dibagi menjadi lima tingkat berdasarkan nilai dihitung. Perbaiki ratio ( *fixed\_ratio* ): rasio tetap dihitung dengan aktiva tetap diving oleh aktiva bersih tidak termasuk

hak minoritas, dibagi menjadi sembilan tingkat berdasarkan nilai dihitung. rasio hutang ( *DSR* ): sesuai dengan nilai dihitung, itu dibagi menjadi 10 tingkatan. Perputaran piutang ( *\_turnover\_piutang* ): perputaran piutang adalah calcu-

lated dengan membagi penjualan bersih dengan piutang rata-rata piutang, dibagi menjadi tujuh tingkat per nilai yang dihitung.

Dampak dari sebuah perusahaan oleh perubahan hukum besar ( *hukum*): tingkat dampak pada perusahaan

yang dibuat oleh perubahan undang-undang utama, dibagi menjadi empat tingkat, dinilai oleh penilai kredit profesional.

indikator makroekonomi ( *makro*): indikator makroekonomi gabungan yang dibuat oleh

lembaga keuangan yang mengukur lingkungan ekonomi saat pinjaman tersebut diterapkan. Kecuali untuk *makro*, variabel lain diskretisasi ke tingkat kategoris dan biner. Itu

distribusi kasus sehubungan dengan masing-masing variabel pada setiap nilai diatur dalam **tabel 3** .

## 4.2 Validasi

Hasil eksperimen model penilaian risiko kredit bawaan jangka pendek yang diusulkan dibandingkan dengan yang dibuat oleh regresi logistik dan Cox model proporsional bahaya untuk mengevaluasi kemampuan klasifikasi model yang diusulkan. Dalam rangka untuk memiliki standar yang sama untuk perbandingan, model yang diusulkan dan regresi logistik menggunakan sampel anced bal-. Cox proportional hazards model yang menggunakan sampel asli. Karena memiliki dua variabel respon, kecuali acara yang akan terjadi atau tidak, perlu untuk mengamati waktu acara dari awal sampai akhir. Memukul hanya mensimulasikan waktu proses secara acak, tidak simul- serentak, menurut dua variabel respon oversampling. Ini mungkin membuat situasi di mana distribusi waktu sampel menyimpang dari dana situasi yang sebenarnya, membiarkan memukul memiliki perbedaan besar dari hasil aplikasi nyata. Di sisi lain, bahaya proporsional Cox model yang dikembangkan untuk menganalisis data yang sangat tidak seimbang, sehingga keseimbangan data tidak langkah yang diperlukan sebelum analisis.

Sebuah model penilaian risiko yang baik harus dapat secara akurat mengidentifikasi kasus default yang dalam waktu singkat setelah menerima pinjaman. Oleh karena itu, makalah ini tidak hanya membandingkan kemampuan diskriminan model yang berbeda tetapi juga membandingkan akurasi klasifikasi keseluruhan model dan penarikan kembali dan presisi kasus bawaan jangka pendek antara model yang berbeda. Tulisan ini digunakan 10 kali lipat cross-validasi untuk menilai kemampuan identifikasi model dalam menghitung rasio akurasi. Dalam 10 kali lipat pendekatan cross-validasi, populasi random domly dibagi menjadi 10 bagian yang sama; 9 bagian digunakan sebagai model pelatihan, dan sisanya 1 dicadangkan untuk pengujian. Operasi ini diulang 10 kali, dan hasil rata-rata adalah hasil akhir dari rasio akurasi. Karena data pengujian tidak berpartisipasi dalam model pelatihan,

## 4.3. analisis variabel

Dalam pohon keputusan, kemampuan diskriminan lebih tinggi dari variabel, yang sebelumnya digunakan untuk cabang. Model penilaian risiko yang diusulkan adalah gabungan dari 11 pohon keputusan. Kita dapat menganalisis frekuensi masing-masing variabel telah menunjukkan dalam 3 tingkat pertama dari 11 pohon keputusan dan ables semacam variabel- berdasarkan tingkat pengaruh mereka pada acara berisiko tinggi (jangka pendek default) rencen occur- peminjam. Hasil eksperimen diatur dalam **tabel 4** . Dari **tabel 4** , Kita dapat melihat bahwa *Latar Belakang* adalah faktor yang paling penting antara lain untuk memprediksi apakah meminjam suatu akan default dalam waktu 1 tahun, karena 64% dari 11 pohon keputusan memilih *Latar Belakang* sebagai variabel pertama untuk cabang di. Ini berarti bahwa reputasi pejabat eksekutif di industri dapat digunakan sebagai indikator penting untuk menentukan apakah default jangka pendek akan terjadi. Ini mungkin karena fakta bahwa kantor eksekutif memiliki reputasi yang baik dalam industri dan sekutu usu- memiliki jaringan interpersonal yang luas dan dengan demikian memiliki sumber yang lebih baik untuk memperoleh dana. Faktor penting kedua adalah *makro*, mewakili lingkungan ekonomi secara keseluruhan. Sekitar 36% dari

**Tabel 1.** Proporsi distribusi kasus default dari masing-masing peringkat variabel.

Variabel	tingkat	Jumlah semua kasus	Proporsi kasus bawaan jangka pendek sampel asli
<i>usia</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>factory_management</i>	1	100	100.00%
	2	100	100.00%
<i>rencana bisnis</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>Latar Belakang</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>bank_evaluation</i>	1	100	100.00%
	2	100	100.00%
<i>competitor_evaluation</i>	1	100	100.00%
	2	100	100.00%
<i>modal</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>liability_ratio</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>fixed_ratio</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>DSR</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>receivable_turnover</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%
<i>hukum</i>	1	100	100.00%
	2	100	100.00%
	3	100	100.00%
	4	100	100.00%

**Tabel 5.** Struktur cabang variabel dalam pohon keputusan.

variabel	lapisan pertama	lapisan kedua	lapisan ketiga
<i>Latar Belakang</i>	• •%	•%	•%
<i>makro</i>	• •%	• •%	•%
<i>liability_ratio</i>	•%	•%	• •%
<i>fixed_ratio</i>	•%	•%	•%
<i>DSR</i>	•%	•%	• •%
<i>competitor_evaluation</i>	•%	•%	•%
<i>hukum</i>	•%	•%	•%
<i>receivable_turnover</i>	•%	•%	•%
<i>modal</i>	•%	•%	• •%
<i>pengelolaan</i>	•%	•%	•%
<i>bank_evaluation</i>	•%	•%	• •%
<i>usia</i>	•%	•%	•%
<i>factory_management</i>	•%	•%	•%

pohon keputusan memilih *makro* sebagai variabel cabang pertama. Ketika bank menyadari menurun kondisi ekonomi, biasanya dimulai untuk memberikan kontrol yang lebih ketat atas permohonan kredit yang dapat mengurangi dana yang tersedia di pasar. Hal ini tidak terduga untuk melihat bahwa peminjam rentan ke default dalam lingkungan ekonomi seperti. Di antara semua 13 variabel, adalah menarik untuk melihat bahwa hanya 2 variabel-variabel ini dipilih sebagai variabel bercabang pertama dalam 11 pohon keputusan.

**Ketujuh variables- latar belakang, makro, liability\_ratio, fixed\_ratio, DSR, competitor\_evaluation, dan hukum-juga** sangat prediktif peristiwa default; variabel-variabel ini terjadi pada frekuensi yang relatif tinggi di lapisan kedua dari pohon keputusan, Total menyumbang 91% dari cabang tingkat kedua dalam 11 pohon keputusan. Variabel ini aremainly digunakan untuk menggambarkan struktur keuangan internal peminjam dan lingkungan yang kompetitif eksternal. Berdasarkan hasil di atas, kita dapat melihat bahwa lingkungan secara keseluruhan ekonomi, aset perusahaan, dan struktur utang adalah faktor yang paling penting untuk menentukan apakah peminjam akan default dalam waktu 1 tahun.

Data pinjaman di variabel makalah ini havemany didefinisikan secara internal oleh insti- tution keuangan sesuai dengan kebutuhan mereka sendiri. Hal ini sangat sulit untuk menemukan informasi keuangan lainnya yang mengandung variabel yang sama, sehingga sulit untuk membuat perbandingan langsung dengan hasil dari penelitian lain. Namun, variabel yang mempengaruhi **kejadian jangka pendek default- makro, liabil- ity\_ratio, fixed\_ratio, dan DSR-yang semua digunakan untuk menilai lingkungan ekonomi** secara keseluruhan atau aset peminjam, struktur utang. Kecenderungan ini sangat mirip dengan model-model penilaian risiko biner yang digunakan variabel keuangan untuk memperkirakan probabilitas default (PD).

#### 4.4. hasil diskriminan

tabel 5 terorganisir hasil percobaan selama tiga model penilaian risiko yang telah diuji. Perkiraan PD dari model Cox digunakan untuk menurunkan tabel 5 dihitung

**Tabel 5.** tingkat recall diskriminan model penilaian risiko.

	Model yang diusulkan		Regresi logistik		Model Cox	
	Penarikan	ketelitian	Penarikan	ketelitian	Penarikan	ketelitian
kasus non standar	• •. •%	• •. •%	• •. •%	• •. •%	• •. •%	• •. •%
Short-term default kasus	• •. •%	• •. •%	• •. •%	• •. •%	• •. •%	• •. •%
Total	• •. •%		• •. •%		• •. •%	



menggunakan  $t = 12$  bulan. Di antara tiga model, model Cox memiliki tertinggi (91,9%) berlebihan semua akurasi diskriminan. Penarikan kembali model yang diusulkan (88,8%) sedikit lebih rendah dari yang model Cox. Kinerja regresi logistik adalah yang terburuk dalam hal berlebihan semua tingkat recall dari model yang diusulkan dan model Cox. Meskipun model Cox memiliki tingkat recall tertinggi secara keseluruhan, mengingat dan presisi dari kasus bawaan jangka pendek yang terburuk di antara tiga model; keduanya di bawah 50%. Hal ini menunjukkan bahwa model Cox tidak dapat secara akurat membedakan kasus kasus pendek termdefault. Perhatikan bahwa karena kemampuan avail- data, model Cox dibangun dan model kinerja dievaluasi dari statis, loandata aplikasi, yaitu,

Makalah ini mengusulkan sebuah model penilaian risiko berbasis pohon keputusan. Meskipun tingkat recall keseluruhan sedikit lebih rendah dari yang model Cox, recall dan presisi yang 81,9% dan 83,3%, masing-masing, dalam hal kasus pendek termdefault diskriminatif. Thismodel, menggunakan pendekatan diskriminan biner, melakukan signifikan lebih baik daripada Coxmodel di identify- ing kasus bawaan jangka pendek. Di sisi lain, regresi logistik juga menggunakan pendekatan criminant dis biner, recall dan presisi pada kasus bawaan jangka pendek (46% dan 64,2%, masing-masing) lebih baik daripada yang oleh model Cox (45,5% dan 45,5%, masing-masing) tapi lebih buruk daripada yang oleh model yang diusulkan (81,9% dan 83,3%, masing-masing).

## 5. Kesimpulan dan arah masa depan

Makalah ini mengusulkan sebuah model penilaian risiko kredit baru untuk mengidentifikasi kasus pendek termdefault dengan pendekatan berbasis pohon keputusan, terintegrasi dengan Bagging dan memukul untuk meningkatkan mengenai efektivitas effec dan stabilitas pohon keputusan. Model ini dapat menyaring kasus bawaan jangka pendek, yang membantu lembaga keuangan untuk memperkirakan potensi kerugian keuangan dan juga untuk menyesuaikan kebijakan kredit. Dibandingkan dengan model penilaian risiko tradisional untuk memperkirakan jika kasus akan default, model yang diusulkan berfokus pada pencegahan kerugian dengan mengidentifikasi kasus-kasus yang mungkin default dalam jangka pendek, sehingga lembaga keuangan mungkin masih dapat membuat keuntungan dari pinjaman default. Di sisi lain, dibandingkan dengan model penilaian risiko kredit lain yang menggunakan konsep analisis survival untuk memperkirakan waktu default, metode yang diusulkan digunakan klasifikasi biner untuk berhasil mengidentifikasi kasus pendek termdefault, orang-orang yang shouldhave pernah addressedwithhigher risiko. Menurut hasil eksperimen, tingkat recall dari model yang diusulkan adalah lebih baik daripada regresi logistik dan model Cox.

Ketika menerapkan model penilaian risiko dalam praktek, lembaga keuangan lebih memilih untuk memiliki aturan klasifikasi yang jelas untuk menerima atau menolak aplikasi pinjaman. Model penilaian risiko pro diajukan dalam makalah ini menggunakan 11 sampel bootstrap untuk membangun 11 pohon keputusan, dimana waktu yang dibutuhkan untuk membangun model yang mirip dengan model yang sudah ada dan tidak terlalu membutuhkan sumber daya yang besar komputasi. Singkatnya, jangka pendek kredit bawaan model penilaian risiko berbasis pohon keputusan yang diusulkan cocok untuk membantu para pengambil keputusan di institusi-institusi keuangan untuk membuat keputusan pinjaman.

Kami telah berhasil menerapkan Bagging dan memukul bersama dengan pohon keputusan untuk mengidentifikasi kasus pendek termdefault yang mencapai sekitar 7% dari total kasus. Satu arah yang menarik untuk penelitian masa depan adalah bekerja pada dataset standar yang sangat rendah dengan sampel mungkin kurang dari 100 kasus default. semacam ini rendah-default sifat portofolio perusahaan sangat umum di kredit korporasi grosir dan besar. Berdasarkan pengalaman sukses kami dalam penelitian ini,

akan menarik untuk melihat efek dari penerapan struktur yang mirip dengan lingkungan tersebut dan untuk membandingkan model standar industri lainnya.

## Ucapan Terima Kasih

Para penulis ingin menyampaikan penghargaan yang tulus kepada pengulas anonim untuk gestions nyarankan- berharga mereka.

## pendanaan

Para penulis mengucapkan terima kasih kepada Departemen Ilmu dan Teknologi Republik China untuk mendukung finansial penelitian ini di bawah nomor kontrak NSC 102-2410-H-145-001 dan PALING 103-2410-H- 145-002.

## Referensi

- Altman, EI (1968). rasio keuangan, analisis diskriminan dan prediksi kebangkrutan perusahaan.  
*J. Keuangan* 23 (4): 589-609.
- Baesens, B., Van Gestel, T., Stepanova, M., Van den Poel, D., Vanthienen, J. (2005). Jaringan syaraf analisis survival untuk data pinjaman pribadi. *J. Oper. Res. Soc.* 56 (9): 1089-1098. Banasik, J., Crook, JN, Thomas, LC (1999). Tidak jika tapi ketika akan peminjam default. *J. Oper. Res. Soc.* 50 (12): 1185-1190.
- Bonfim, D. (2009). driver risiko kredit: Mengevaluasi kontribusi informasi tingkat perusahaan dan dinamika ekonomi makro. *J. Perbankan Keuangan* 33 (2): 281-299. Breiman, L. (1996). prediktor mengantongi. *Mach. Belajar.* 24 (2): 123-140.
- Bank Sentral Republik Cina. (2011). Kondisi dan Kinerja Bank Domestik, Keempat Kuartal 2011, Taiwan.
- Chawla, NV, Bowyer, KW, Hall, LO, Kegelmeyer, WP (2002). Memukul: minoritas sintetis berlebihan teknik sampling. *J. artif. Intell. Res.* 16: 321-357. Cox, DR (1972). model regresi dan kehidupan-tabel. *J. Kerajaan Stat. Soc. Ser. B* 34 (2): 187-220. Duffie, D., Saita, L., Wang, K. (2007). Multi-periode standar perusahaan prediksi dengan covari- stokastik ates. *J. Econ Keuangan.* 83 (3): 635-665.
- Eklund, T., Larsen, K., Berghardsen, E. (2001). Model untuk menganalisis risiko kredit di sektor usaha. *Econ. Banteng.* Q3 (01): 99-106.
- Mester, LJ (1997). Apa gunanya credit scoring? Federal Reserve Bank of Philadelphia Bisnis Review, September / Oktober, 3-16.
- Narain, B. (1992). analisis kelangsungan hidup dan keputusan pemberian kredit. Dalam: Thomas, LC, Crook, JN, Edelman, DB, eds. *Scoring Kredit dan Pengendalian Kredit* (pp. 109-121). Oxford, UK: OUP. Quinlan, JR (1986). Induksi pohon keputusan. *Mach. Belajar.* 1 (1): 81-106. Quinlan, JR (1993). *C4.5: Program untuk Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers.
- Roszbach, K. (2004). Bank pinjaman kebijakan, penilaian kredit, dan kelangsungan hidup pinjaman. *Rev. Econ. Stat.* 86 (4): 946-958.
- Shumway, T. (2001). Peramalan bankruptcy more akurat: a hazard model sederhana. *J. Bus.* 74 (1): 101-124.
- Thomas, LC (2000). Sebuah survei dari kredit dan scoring perilaku: peramalan risiko keuangan pinjaman kepada konsumen. *Int. J. Peramalan* 16 (2): 149-172.
- Tong, ENC, Mues, C., Thomas, LC (2012). model campuran obat di credit scoring: jika dan ketika bor- pendayung default. *Eur. J. Oper. Res.* 218 (1): 132-139.