



Implementasi SMOTE dan *Under Sampling* pada *Imbalanced Dataset* untuk Prediksi Kebangkrutan Perusahaan

Wilda Imama Sabilla¹, Candra Bella Vista²

¹Politeknik Negeri Malang, Jurusan Teknologi Informasi, email: wildaimama@polinema.ac.id

²Politeknik Negeri Malang, Jurusan Teknologi Informasi, email: bellavista@polinema.ac.id

[1] Abstrak

Kebangkrutan pada suatu perusahaan menjadi masalah yang serius karena dapat menyebabkan kerusakan ekonomi serta konsekuensi sosial lainnya. Sangat penting untuk melakukan prediksi kebangkrutan sedini mungkin karena prediksi ini dapat bermanfaat untuk evaluasi serta merencanakan tindakan pencegahan dalam menghindari kebangkrutan. Prediksi kebangkrutan merupakan salah satu permasalahan imbalanced classification karena data dengan kelas bangkrut jauh lebih sedikit daripada kelas tidak bangkrut. Penelitian ini bertujuan untuk menghasilkan model klasifikasi yang baik untuk melakukan prediksi kebangkrutan. Resampling diterapkan pada data latih agar menghasilkan model klasifikasi yang lebih optimal. Metode resampling yang digunakan adalah kombinasi SMOTE dan under sampling. Metode klasifikasi yang digunakan untuk prediksi adalah multilayer perceptron dan complement naïve bayes. Performa prediksi dihitung menggunakan skor recall, ROC AUC, dan PR AUC. Berdasarkan hasil pengujian, penggunaan SMOTE dan under sampling cukup signifikan dalam memperbaiki model klasifikasi pada multilayer perceptron. Pada prediksi menggunakan complement naïve bayes, nilai recall dan PR AUC juga meningkat. Recall terbaik yang diperoleh sebesar 95,45% dengan metode complement naïve bayes. Untuk ROC AUC tertinggi dengan resampling juga diperoleh menggunakan complement naïve bayes sebesar 87,80%. Oleh karena itu, disimpulkan bahwa prediksi kebangkrutan menggunakan teknik resampling yaitu SMOTE dan under sampling dapat menghasilkan performa baik untuk pendeteksian kelas bangkrut.

Kata kunci: *imbalanced dataset, prediksi kebangkrutan, smote, under sampling, multilayer perceptron, complement naïve bayes*

[2] Abstract

Company bankruptcy becomes a serious problem because it can cause economic damage and other social consequences. It's very important to predict bankruptcy as early as possible because prediction can be useful for evaluation and planning to avoid bankruptcy. Bankruptcy prediction is one of the imbalanced classification problems because the data with the bankrupt class is far less than the non-bankrupt class. This study aims to produce a good classification model for predicting bankruptcy. Resampling used a combination of SMOTE and under sampling, is applied to the training data to produce more optimal classification model. The classification method used for prediction is multilayer perceptron and complement naïve bayes. Predictive performance was calculated using recall, ROC AUC, and PR AUC. Based on the test, using SMOTE and under

sampling is quite significant in improving the classification model on the multilayer perceptron. Resampling in complement naïve bayes also increased. recall and PRAUC scores The best recall obtained was 95.45% with the complement naïve bayes method. The highest ROC AUC with resampling was also obtained using complement naïve bayes of 87.80%. Therefore, it's concluded that bankruptcy prediction using resampling with SMOTE and under sampling, can produce good performance for detecting bankruptcy.

Keywords: *imbalanced dataset, bankruptcy prediction, smote, under sampling, multilayer perceptron, complement naïve bayes*

1. Pendahuluan

Kebangkrutan disebut juga likuidasi perusahaan atau penutupan perusahaan atau insolvensi, merupakan kondisi terjadinya kegagalan keuangan dan kegagalan ekonomi. Sebelum terjadi kebangkrutan, biasanya perusahaan mengalami *financial distress* (kesulitan keuangan) yang ditandai dengan penurunan kondisi keuangan [1]. Kebangkrutan perusahaan menjadi masalah yang serius pada pasar finansial karena dapat menyebabkan kerusakan ekonomi serta konsekuensi sosial lainnya [2]. Kebangkrutan perusahaan akan berdampak besar pada stakeholder perusahaan. Untuk perusahaan *go public*, kebangkrutan berdampak pada dihapusnya saham perusahaan pada bursa efek. Seperti yang terjadi pada 20 perusahaan yang dihapus dari Bursa Efek Indonesia pada tahun 2011 hingga 2015 [3].

Melihat dampak yang terjadi akibat kebangkrutan perusahaan maka prediksi kebangkrutan menjadi suatu pekerjaan penting bagi institusi keuangan [4]. Prediksi kebangkrutan bertujuan untuk membuat model prediksi yang efektif guna mengetahui resiko kebangkrutan berdasarkan data-data keuangan yang dimiliki oleh perusahaan. Prediksi kebangkrutan sedini mungkin dapat bermanfaat bagi perusahaan untuk melakukan evaluasi serta merencanakan tindakan pencegahan agar kebangkrutan dapat dihindari [3].

Sebagian besar data di dunia nyata merupakan kumpulan data yang tidak seimbang (*imbalanced dataset*). *Imbalanced dataset* memiliki perbedaan jumlah data yang cukup besar antara kelas satu dengan kelas lainnya. Kelas yang memiliki data sampel sangat besar disebut kelas mayoritas (*majority class*), sedangkan kelas data yang memiliki sampel sedikit disebut kelas minoritas (*minority class*) [5]. Data dari kelas minoritas seringkali mengalami kesalahan pelabelan saat diproses pada sebuah model pengklasifikasi (*classifier*). Oleh karena itu dibutuhkan teknik tertentu untuk pengambilan ulang data sampel (*resampling*) seperti *undersampling* atau *oversampling* untuk mengatasi permasalahan tersebut [6].

Permasalahan *imbalanced dataset* banyak ditemui di berbagai bidang. Di antaranya pada bidang kesehatan contohnya pada prediksi epilepsi dan Parkinson dan penyakit *liver* [7]. Di bidang ekonomi dan bisnis manajemen, *imbalanced dataset* yang sering ditemui antara lain data untuk prediksi kebangkrutan perusahaan dan deteksi penipuan (*fraud detection*) kartu kredit [8] dan asuransi [9].

Teknik *resampling* yang digunakan pada penelitian terkait prediksi kebangkrutan antara lain menggunakan *oversampling* [5], *cluster-based boosting* [6], dan *hybrid under-sampling* [10]. Sedangkan pada penelitian sebelumnya, berbagai algoritma klasifikasi digunakan untuk memprediksi kebangkrutan. Algoritma klasifikasi yang dipakai antara lain *Bayesian Network* [2], *Support Vector Machine* dan *Artificial Neural Network* [11].

Penelitian ini memprediksi kebangkrutan perusahaan menggunakan dua metode pengklasifikasi yaitu *Multilayer Perceptron* dan *Complement Naïve Bayes*. Atribut dari data keuangan perusahaan akan dipilih berdasarkan nilai korelasi atribut dengan kelas. Karena data memiliki

jumlah kelas yang tidak seimbang, maka akan dilakukan *resampling* menggunakan SMOTE dan *undersampling*. Performa metode pengklasifikasi akan diukur menggunakan *Recall*, *ROC AUC* dan *PR AUC*. Dari metrik pengukuran tersebut akan dianalisis apakah *SMOTE* dan *under sampling* berpengaruh terhadap performa metode pengklasifikasi.

2. Tinjauan Pustaka

Bagian ini menjelaskan mengenai landasan teori yang digunakan sebagai acuan dalam melaksanakan metode penelitian.

2.1 Pearson Correlation Coefficient

Korelasi (*correlation*) digunakan untuk mengukur nilai hubungan antara dua variabel. Dengan menggunakan *Pearson Correlation Coefficient*, maka akan dihasilkan nilai korelasi linear dalam rentang -1 dan 1. Korelasi yang baik memiliki nilai mendekati 1 atau -1. Nilai -1 berarti korelasi negatif sempurna terjadi ketika nilai variabel meningkat ketika variabel lainnya menurun, sedangkan nilai 1 menandakan korelasi positif sempurna. Nilai korelasi 0 menunjukkan bahwa kedua variabel tidak memiliki korelasi [12]. Persamaan 1 merupakan persamaan dari *Pearson Correlation Coefficient*.

$$r_{xy} = \frac{n(\sum xy) - \sum x \sum y}{\sqrt{[\sum x^2 - (\sum x)^2]} \sqrt{[n \sum y^2 - (\sum y)^2]}} \quad (1)$$

Dimana:

- r_{xy} : Pearson's Product Moment Correlation Coefficient
- n : banyaknya pasangan data x dan y
- $\sum x$: total jumlah nilai variabel x
- $\sum y$: total jumlah nilai variabel y
- $\sum x^2$: kuadrat dari total jumlah nilai variabel x
- $\sum y^2$: kuadrat dari total jumlah nilai variabel y
- $\sum xy$: total jumlah dari perkalian antara variabel x dan y

2.2 Synthetic Minority Oversampling Technique (SMOTE)

Klasifikasi pada *imbalanced dataset* cukup sulit dilakukan karena terlalu sedikit data dari kelas minoritas untuk dapat menghasilkan model klasifikasi yang efektif. Untuk mengatasi hal tersebut, salah satu cara yang bisa dilakukan adalah melakukan *oversampling* yaitu melakukan duplikasi data dari kelas minoritas pada data latih agar menghasilkan model yang lebih baik.

Synthetic Minority Oversampling Technique (SMOTE) memilih data pada kelas minoritas secara acak kemudian mencari k data kelas minoritas terdekat dari data tersebut [13]. Dari k data kelas minoritas tersebut, akan dipilih satu data secara acak yang selanjutnya dihubungkan dengan data awal yang dipilih untuk membentuk segmen garis pada ruang fitur. Data sintesis dihasilkan menggunakan kombinasi convex antara dua data yang dipilih secara acak tadi. Pendekatan ini dinilai cukup efektif karena data sintesis yang dibentuk relatif dekat dalam ruang fitur dengan data yang ada dari kelas minoritas. Prosedur ini dapat digunakan untuk membuat sebanyak mungkin data sintetik untuk kelas minoritas yang diperlukan. Namun agar menghasilkan data yang lebih baik, teknik SMOTE menyarankan penggunaan *under-sampling* acak untuk menyeimbangkan distribusi kelas. Kekurangan dari SMOTE adalah data sintesis yang dibentuk tidak memperhatikan kelas mayoritas sehingga memungkinkan adanya tumpang tindih data.

2.3 Random Under Sampling

Random under sampling melakukan pemilihan data secara acak dari kelas mayoritas untuk dihapus dari kumpulan data latih. Dengan menjalankan *random under sampling*, maka data latih

dari kelas mayoritas akan berkurang jumlahnya. Proses *under sampling* dapat diulang hingga diperoleh distribusi kelas yang diinginkan pada data latih. Pendekatan ini dapat diterapkan pada kumpulan data dengan kelas yang tidak seimbang dimana kelas minoritas cukup untuk pembuatan model. Kekurangan dari *undersampling* adalah data dari kelas mayoritas yang dihapus adalah data acak sehingga ada kemungkinan data tersebut adalah data yang berguna atau bahkan sangat penting dalam pembangunan model klasifikasi yang baik [14]. Pada *under sampling*, dimungkinkan dibuat jumlah data yang sama dari kelas mayoritas dan minoritas atau hanya mengurangi data mayoritas hingga jumlah tertentu.

2.4 Multilayer Perceptron (MLP)

Multilayer perceptron adalah salah satu model jaringan syaraf tiruan (*artificial neural network*). *Multilayer perceptron* setidaknya memiliki tiga layer yaitu *input layer*, *hidden layer*, dan *output layer*. Metode ini melakukan *update* bobot pada setiap layer pada saat membangun model klasifikasi menggunakan *backpropagation*.

Berikut ini adalah langkah-langkah pada metode *multilayer perceptron* [15]:

1. Menentukan bobot awal dengan bilangan acak kecil.
2. Menghitung *output* aktual dari *neuron* pada *hidden layer* dan *output layer* menggunakan fungsi aktivasi.
3. Untuk setiap data latih dilakukan proses pelatihan untuk pembuatan model dengan menggunakan algoritma *backpropagation*.
4. Proses *update* bobot dimulai dari *output layer* menuju ke *hidden layer*. Hasil dari penghitungan bobot baru pada *output layer* akan digunakan untuk menghitung bobot baru pada *hidden layer*.

2.5 Complement Naïve Bayes (CNB)

Complement Naïve Bayes adalah adaptasi dari standar *Multinomial Naïve Bayes* yang sesuai untuk kumpulan data tidak seimbang (*imbalance dataset*). *Complement Naïve Bayes* menggunakan statistik dari komplemen setiap kelas untuk menghitung bobot model pengklasifikasi. Penelitian mengenai *Complement Naïve Bayes* menunjukkan bahwa estimasi parameter pada *Complement Naïve Bayes* lebih stabil dibandingkan dengan *Multinomial Naïve Bayes* [16]. Terutama pada permasalahan klasifikasi teks, *Complement Naïve Bayes* memiliki performa dengan margin yang cukup besar terhadap *Multinomial Naïve Bayes*.

Complement Naïve Bayes melakukan perhitungan bobot sebagaimana pada persamaan 2, 3 dan 4 yang merupakan penjumlahan dari semua dokumen j yang tidak terdapat di kelas c .

$$\hat{\theta}_{ci} = \frac{\alpha_i + \sum_{j:y_j \neq c} d_{ij}}{\alpha + \sum_{j:y_j \neq c} \sum_{k=1}^K d_{kj}} \quad (2)$$

$$w_{ci} = \log \hat{\theta}_{ci} \quad (3)$$

$$w_{ci} = \frac{w_{ci}}{\sum_j |w_{cj}|} \quad (4)$$

Dimana:

d_{ij} : jumlah atau nilai tf-idf dari term i di dokumen j

α_i : *smoothing hyperparameter* seperti yang terdapat pada *Multinomial Naïve Bayes*

α : penjumlahan semua nilai α_i

Sedangkan untuk aturan klasifikasinya ditunjukkan pada Persamaan 5.

$$\hat{c} = \arg \min_c \sum_t t_i w_{ct} \quad (5)$$

dimana suatu dokumen dilabelkan ke kelas yang memiliki kecocokan komplemen terendah.

2.6 Model Evaluasi

Pada klasifikasi dengan *imbalanced dataset*, kelas mayoritas seringkali dilabeli sebagai kelas negatif sedangkan kelas minoritas dianggap sebagai kelas positif. Untuk melihat kesesuaian klasifikasi biasa digunakan matriks yang disebut dengan *confusion matrix*. Selain dapat digunakan untuk melihat kinerja model, *confusion matrix* juga dapat menunjukkan seberapa banyak kesalahan prediksi baik pada kelas minoritas maupun kelas mayoritas [14] .

Untuk klasifikasi dengan dua kelas (*binary classification*) dengan kelas negatif (0) dan kelas positif (1), maka *confusion matrix* dapat dilihat di Tabel 1.

Tabel 1. *Confusion Matrix*

	Prediksi Positif	Prediksi Negatif
Kelas Positif	True Positive (TP)	False Negative (FN)
Kelas Negatif	False Positive (FP)	True Negative (TN)

Recall adalah metrik yang mengukur jumlah prediksi positif benar (TP) dari semua prediksi positif yang dihasilkan. *Recall* dapat menunjukkan indikasi prediksi positif yang terlewatkan (FN). Persamaan untuk menghitung *recall* terdapat pada Persamaan 6.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

Untuk klasifikasi pada *imbalanced dataset*, *recall* digunakan untuk mengukur kebenaran prediksi dari kelas minoritas. Nilai *recall* berada pada rentang 0 sampai 1 dimana 1 menunjukkan hasil *recall* sempurna.

AUC adalah *Area Under the Curve*, ROC AUC mendefinisikan area di bawah kurva ROC. Kurva ROC (*Receiver Operating Characteristic*) merupakan alat yang digunakan untuk mengukur hasil prediksi pada data dengan dua kelas (*binary classification*). ROC membuat plot antara *False Positive Rate* (sumbu x) dan *True Positive Rate* (sumbu y) dengan nilai ambang antara 0.0 dan 1.0. Rumus untuk *True Positive Rate* dan *False Positive Rate* terdapat pada Persamaan 7 dan 8.

$$True\ Positive\ Rate = Recall = \frac{TP}{TP + FN} \quad (7)$$

$$False\ Positive\ Rate = \frac{FP}{FP + TN} \quad (8)$$

ROC AUC dapat digunakan untuk mengetahui ketepatan model klasifikasi. Skor ROC AUC sama dengan menghitung korelasi peringkat antara prediksi dan target.

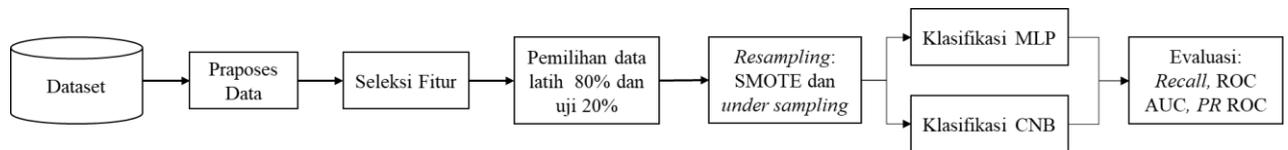
PR AUC mendefinisikan area di bawah kurva PR. Kurva PR mengkombinasikan *precision* dan *recall* (TPR). Untuk setiap ambang batas dihitung *precision* dan *recall* serta diplot ke kurva PR dimana sumbu x adalah *recall* dan sumbu y adalah *precision*. Rumus perhitungan *precision* terdapat pada Persamaan 9.

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

Sama halnya dengan ROC AUC, PR AUC juga dapat digunakan untuk mengukur performa model klasifikasi terutama pada *binary classification*. PR AUC cocok untuk *imbalanced dataset* dimana keberadaan kelas positif (kelas minoritas) lebih diutamakan daripada kelas negatif (kelas mayoritas) [17].

3. Metode Penelitian

Pada bagian ini akan dijelaskan metodologi penelitian. Arsitektur sistem pada penelitian ini dapat dilihat pada Gambar 1. Dataset dilakukan *preprocessing* (praproses) data terlebih dahulu untuk menghasilkan data dengan rentang seragam. Selanjutnya dilakukan seleksi atribut berdasarkan korelasi dengan atribut kelas. Sebelum dilakukan pembangunan model klasifikasi, diterapkan SMOTE dan *under sampling* pada data latih untuk menangani permasalahan kelas data yang tidak seimbang. Setelah terbentuk model klasifikasi baik dengan MLP dan CNB, dilakukan prediksi data latih. Hasil prediksi diuji menggunakan *recall*, ROC AUC, serta PR AUC.



Gambar 1. Arsitektur Sistem

3.1 Dataset

Dataset yang digunakan di penelitian ini merujuk pada penelitian berikut [4] yang dikumpulkan oleh Taiwan Economic Journal pada tahun 1999–2009. Status kebangkrutan ditentukan oleh peraturan bisnis yang ditetapkan oleh Bursa Efek Taiwan. Perusahaan yang diambil datanya setidaknya telah melengkapi informasi publik selama 3 tahun sebelum terjadinya krisis finansial. Perusahaan tersebut juga harus dapat dibandingkan dengan perusahaan lain dengan skala industri yang sama. Bidang perusahaan pada data ini terdiri dari 346 industri manufaktur, 39 industri layanan (pengiriman, wisata, ritel), dan 93 bidang lainnya.

Data memiliki 96 atribut dimana salah satu atributnya adalah label kelas. Atribut kelas bernama “Bankrupt?” yang bernilai 1 untuk kelas bangkrut dan 0 untuk kelas tidak bangkrut. Kelas bangkrut sebagai kelas minoritas memiliki 220 data dan kelas tidak bangkrut sebagai kelas mayoritas memiliki 6599 data.

3.2 Praproses Data

Praproses bertujuan untuk menyeragamkan rentang data antara 0 sampai 1. Sebelumnya atribut “net income flag” dihapus dari daftar atribut karena semua baris bernilai 1. Agar nilai atribut antara 0 dan 1, maka setiap nilai atribut dibagi dengan nilai maksimal atribut dari semua data. Sebagai contoh untuk atribut “Operating Expense Rate” nilai maksimal atribut tersebut adalah 9,990,000,000. Maka untuk data dengan nilai atribut 9,000,000,000 hasil praprosesnya menjadi $9,000,000,000 / 9,990,000,000 = 0.900900901$.

3.2 Seleksi Fitur

Karena jumlah fitur cukup banyak yaitu 94 fitur, maka perlu dilakukan pemilihan fitur untuk mengurangi jumlah fitur dan menghasilkan fitur yang bisa memperbaiki kinerja algoritma klasifikasi. Pemilihan dilakukan dengan menghitung korelasi antara fitur dan kelas menggunakan *Pearson Correlation Coefficient*. Dari 94 fitur dipilih 32 fitur yang nilai korelasinya > 0.1 , baik korelasi positif ataupun korelasi negatif. Tabel 2 menunjukkan fitur yang terpilih beserta nilai korelasinya terhadap atribut kelas.

Tabel 2. Fitur terpilih dan nilai korelasinya

No	Nama Fitur	Nilai Korelasi	No	Nama Fitur	Nilai Korelasi
1	ROA(C) before interest and depreciation before interest	-0.2608	17	Working Capital to Total Assets	-0.1931
2	ROA(A) before interest and % after tax	-0.2829	18	Cash/Total Assets	-0.1001
3	ROA(B) before interest and depreciation after tax	-0.2730	19	Current Liability to Assets	0.1945
4	Operating Gross Margin	-0.1000	20	Working Capital/Equity	-0.1472
5	Tax rate (A)	-0.1097	21	Current Liabilities/Equity	0.1538
6	Net Value Per Share (B)	-0.1654	22	Retained Earnings to Total Assets	-0.2178
7	Net Value Per Share (A)	-0.1655	23	Total expense/Assets	0.1390
8	Net Value Per Share (C)	-0.1648	24	Current Liabilities/Equity	0.1538
9	Persistent EPS in the Last Four Seasons	-0.2196	25	Equity to Long-term Liability	0.1390
10	Operating Profit Per Share (Yuan ¥)	-0.1420	26	CFO to Assets	-0.1154
11	Per Share Net profit before tax (Yuan ¥)	-0.2014	27	Current Liability to Current Assets	0.1713
12	Debt ratio %	0.2502	28	Liability-Assets Flag	0.1392
13	Net worth/Assets	-0.2502	29	Net Income to Total Assets	-0.3155
14	Borrowing dependency	0.1765	30	Gross Profit to Sales	-0.1000
15	Operating profit/Paid-in capital	-0.1411	31	Net Income to Stockholder's Equity	-0.1810
16	Net profit before tax/Paid-in capital	-0.2079	32	Liability to Equity	0.1668

3.3 Pemilihan Data Latih dan Data Uji

Untuk dapat dilakukan prediksi, dataset harus dibagi terlebih dahulu ke dalam data latih dan data uji. Pada penelitian ini keseluruhan data dibagi menjadi 80%:20% untuk 80% data latih dan 20% data uji. Data dipilih secara acak dengan perbandingan yang sama antara kelas bangkrut dan tidak bangkrut (*stratify random sampling*). Tabel 3 menunjukkan jumlah data latih dan data uji.

Tabel 3. Pembagian data latih dan data uji

Jenis Data	Jumlah Data		
	Kelas 1	Kelas 0	Total
Data latih	176	5279	5455
Data uji	44	1320	1364

3.4 Resampling dengan SMOTE dan under sampling

Karena dataset yang digunakan merupakan *imbalanced dataset*, maka dari data latih yang telah dipilih akan dilakukan *over sampling* dengan SMOTE dan dilanjutkan dengan *under sampling*.

Pada penelitian ini, digunakan *sampling_strategy* 0.1 pada SMOTE dan 0.5 untuk *under sampling*. Jumlah data hasil *resampling* dapat dilihat pada Tabel 4.

Tabel 4. Pembagian data latih dan data uji

Label Kelas	Jumlah data latih
1	1054
0	527

3.5 Prediksi dengan *Multilayer Perceptron* dan *Complement Naïve Bayes*

Data latih yang telah dipilih akan digunakan untuk membentuk model klasifikasi *multilayer perceptron* dan *complement naïve bayes*. Pada prediksi dengan *multilayer perceptron*, model dibangun dengan satu *hidden layer* yang memiliki 10 buah *node*. Sedangkan pada *complement naïve bayes* digunakan *smoothing* parameter dengan alpha bernilai 1. Selanjutnya dilakukan prediksi label kelas dari data uji menggunakan model yang telah dibangun.

3.6 Evaluasi Hasil Prediksi

Setelah diperoleh hasil prediksi dari data uji, maka akan dibentuk *confusion matrix* untuk mengetahui TP, TN FP, dan FN. Selanjutnya digunakan metric pengukuran berupa *recall*, ROC AUC, dan PR AUC. Ketiga metrik ini dipilih karena dapat memperlihatkan kinerja model klasifikasi pada *dataset imbalanced*. *Recall* yang utamanya dapat melihat performa model klasifikasi dalam pendeteksian kelas positif sangat diperlukan pada prediksi kebangkrutan. Hal ini dikarenakan prediksi kelas dengan label 1 (bangkrut) dibutuhkan perusahaan atau lembaga keuangan untuk merencanakan tindakan agar kebangkrutan dapat dicegah.

4. Hasil dan Pembahasan

Bagian ini menjelaskan hasil uji coba pada penelitian ini serta analisis hasil uji coba.

4.1 Hasil Uji Coba Prediksi dengan *Multilayer Perceptron*

Uji coba dilakukan dengan menjalankan klasifikasi dengan *multilayer perceptron* sebanyak 10 kali. Setiap kali percobaan akan dicatat nilai *recall*, ROC AUC, dan PR AUC. Percobaan dilakukan menggunakan data latih asli dan data latih setelah diterapkan SMOTE dan *under sampling*.

Tabel 5 menunjukkan performa terbaik dari beberapa percobaan prediksi menggunakan *multilayer perceptron*. Dengan menggunakan *resampling* maka performa *recall*, ROC AUC, PR AUC meningkat dengan signifikan. Nilai *recall* naik dari 13.64% menjadi 81.82%. ROC AUC juga meningkat cukup tinggi dari 56.63% ke angka 87.20%. Walaupun perubahannya tidak sebanyak yang lain, tapi pada PR AUC performanya juga meningkat dari 35.48% menjadi 54.64%.

Tabel 5. Performa Prediksi dengan MLP

Metrik Pengukuran	Data latih asli	Data latih setelah <i>resampling</i>
<i>Recall</i>	0,1364	0,8182
ROC AUC	0,5663	0,8720
PR AUC	0,3548	0,5464

4.2 Hasil Uji Coba Prediksi dengan *Complement Naïve Bayes*

Sama halnya dengan percobaan dengan *multilayer perceptron*, percobaan klasifikasi menggunakan *complement naïve bayes* dijalankan sebanyak 10 kali. Nilai *recall*, ROC AUC, dan PR AUC diambil tiap percobaan untuk menghitung performa metode *complement naïve bayes*. Percobaan dijalankan memakai dua jenis data latih yaitu data latih asli dan data latih setelah diterapkan SMOTE dan *under sampling*.

Performa terbaik dari beberapa percobaan prediksi menggunakan *complement naïve bayes* ditunjukkan pada Tabel 6. Nilai *recall* yang diperoleh cukup tinggi baik pada data asli maupun data *resampling* yaitu 93,18% dan 95,45%. Walaupun tidak berbeda jauh, tetapi penggunaan SMOTE dan *under sampling* menurunkan skor ROC AUC dari 88,41% ke angka 87,80%. Sedangkan pada PR AUC kenaikannya cukup kecil yaitu 56,68% menjadi 54,71% dengan *resampling*.

Tabel 6. Performa Prediksi dengan CNB

Metrik Pengukuran	Data latih asli	Data latih setelah <i>resampling</i>
<i>Recall</i>	0,9318	0,9545
ROC AUC	0,8841	0,8780
PR AUC	0,5468	0,5471

4.3 Pembahasan Hasil Uji Coba

Uji coba telah dilakukan menggunakan data latih asli dan data latih setelah SMOTE dan *under-sampling*. Uji coba dilakukan untuk meninjau performa prediksi kebangkrutan dengan *dataset imbalanced* menggunakan *multilayer perceptron* dan *complement naïve bayes*.

Performa terbaik yang ditunjukkan oleh metode klasifikasi MLP berupa skor *recall*, ROC AUC, dan PR AUC masing-masing 81,82%, 87,20% dan 54,64%. Kenaikan performa terjadi cukup signifikan pada metode MLP yang dilakukan *resampling* dibandingkan dengan data asli. Kenaikan yang terjadi untuk *recall*, ROC AUC, dan PR AUC berurutan yaitu 68,18%, 30,57% dan 19,16%. Hal ini menunjukkan bahwa dalam prediksi menggunakan metode MLP, SMOTE serta *under sampling* berpengaruh dalam pembuatan model klasifikasi. Berdasarkan skor dari tiga metric tersebut diketahui model klasifikasi yang dihasilkan lebih baik setelah *resampling*.

Untuk prediksi menggunakan metode CNB nilai *recall*, ROC AUC, dan PR AUC tidak berbeda signifikan antara data latih asli dan data latih setelah *resampling*. Hal ini dimungkinkan karena CNB telah dirancang untuk membuat model klasifikasi yang cukup baik untuk *dataset imbalanced*. Walaupun demikian tetap terjadi peningkatan pada skor *recall* dan PR AUC sebesar 2,27% dan 0,03%.

Secara garis besar, penggunaan SMOTE dan *under sampling* berdampak baik pada performa MLP maupun CNB pada nilai *recall* dan PR AUC. *Recall* dan ROC AUC yang diperoleh juga relatif tinggi di atas 80%. Sedangkan pada PR AUC skor yang diperoleh lebih rendah dimungkinkan karena walaupun model baik mendeteksi kelas positif namun masih kurang baik dalam mendeteksi kelas negatif. *Recall* tertinggi didapat menggunakan CNB dengan *resampling* yaitu 95,45%. Untuk ROC AUC tertinggi dengan *resampling* juga diperoleh menggunakan CNB sebesar 87,80%.

5. Kesimpulan

Prediksi kebangkrutan merupakan salah satu permasalahan *imbalanced classification* karena data dengan kelas bangkrut jauh lebih sedikit daripada kelas tidak bangkrut. Sangat dibutuhkan model klasifikasi yang dapat melakukan klasifikasi dengan baik terutama untuk kelas minoritas yaitu

data dengan kelas bangkrut. Penelitian ini menggunakan SMOTE dan *under sampling* untuk menangani jumlah kelas data yang tidak seimbang pada saat klasifikasi. Metode klasifikasi yang digunakan adalah *multilayer perceptron* dan *complement naïve bayes*. Dari performa yang diperoleh menggunakan skor *recall*, ROC AUC, dan PR AUC penggunaan SMOTE dan *under sampling* cukup signifikan dalam memperbaiki model klasifikasi pada *multilayer perceptron*. Sedangkan untuk metode klasifikasi dengan *complement naïve bayes*, nilai *recall* dan PR AUC juga meningkat. *Recall* terbaik yang diperoleh sebesar 95,45% dengan CNB dan *resampling*. Untuk ROC AUC tertinggi dengan *resampling* juga diperoleh menggunakan CNB sebesar 87,80%. Sehingga dapat disimpulkan penelitian untuk prediksi kebangkrutan menggunakan teknik *resampling* SMOTE dan *under sampling* menghasilkan performa baik untuk pendeteksian kelas bangkrut.

6. Daftar Pustaka

- [1] A. S. Ramadhani and N. Lukviarman, "Perbandingan Analisis Prediksi Kebangkrutan Menggunakan Model Altman Pertama, Altman Revisi, Dan Altman Modifikasi Dengan Ukuran Dan Umur Perusahaan Sebagai Variabel Penjelas (Studi Pada Perusahaan Manufaktur Yang Terdaftar Di Bursa Efek Indonesia)," *Jurnal Siasat Bisnis*, vol. 13, no. 1, pp. 15-28, April 2009.
- [2] Y. Cao, X. Liu, J. Zhai and S. Hua, "A two-stage Bayesian network model for corporate bankruptcy prediction," *International Journal of Finance & Economics*, pp. 1-18, 2020.
- [3] N. H. Maturungan, B. Purwanto and A. K. Irwanto, "Manufacturing Company Bankruptcy Prediction in Indonesia with Altman Z-Score Model," *Journal of Applied Management*, vol. 15, no. 1, pp. 18-24, March 2017.
- [4] D. Liang, C.-C. Lu, C.-F. Tsai and G.-A. Shih, "Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study," *European Journal of Operational Research*, vol. 252, pp. 561-572, 2016.
- [5] T. Le, M. Y. Lee, J. R. Park and S. W. Baik, "Oversampling Techniques for Bankruptcy Prediction: Novel Features from a Transaction Dataset," *Symmetry*, vol. 10, pp. 79-91, 2018.
- [6] T. Le, L. H. Son, M. T. Vo, M. Y. Lee and S. W. Baik, "A Cluster-Based Boosting Algorithm for Bankruptcy Prediction in a Highly Imbalanced Dataset," *Symmetry*, vol. 10, pp. 250-262, 2018.
- [7] P. Vuttipittayamongkol and E. Elyan, "Improved Overlap-based Undersampling for Imbalanced Dataset Classification with Application to Epilepsy and Parkinson's Disease," *International Journal of Neural Systems*, pp. 1-16, 2020.
- [8] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M.-S. Hacid and H. Zeineddine, "An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection," *IEEE Access*, pp. 93010-93022, 2019.
- [9] A. K. I. Hassan and A. Abraham, "Modeling Insurance Fraud Detection Using Imbalanced Data Classification," *Advances in Nature and Biologically Inspired Computing, Advances in Intelligent Systems and Computing*, pp. 117-127, 2016.
- [10] T. Kim and H. Ahn, "A Hybrid Under-sampling Approach for Better Bankruptcy Prediction," *Journal of Intelligence and Information Systems*, pp. 173-190, 2015.
- [11] J. Horak, J. Vrbka and P. Suler, "Support Vector Machine Methods and Artificial Neural Networks Used for the Development of Bankruptcy Prediction Models and their Comparison," *Journal of Risk and Financial Management*, p. 60, 2020.
- [12] Y. Bae and H. Lee, "Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 12, pp. 2521-2535, 2012.

- [13] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal Of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002.
- [14] H. He and Y. Ma, *Imbalanced Learning: Foundations, Algorithms, and Applications* 1st Edition, New Jersey: IEEE Press Wiley, 2013.
- [15] E. P. Cynthia and E. Ismanto, "Jaringan Syaraf Tiruan Algoritma Backpropagation dalam Memprediksi Ketersediaan Komoditi Pangan Provinsi Riau," *Rabit : Jurnal Teknologi Dan Sistem Informasi Univrab*, vol. 2, no. 2, pp. 83-98, 2017.
- [16] J. D. Rennie, L. Shih, J. Teevan and D. R. Karger, "Tackling the poor assumptions of naive bayes text classifiers," *CML*, vol. 3, pp. 616-623, 2003.
- [17] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS ONE*, vol. 10, no. 3, p. e0118432, 2015.