

**PENERAPAN SELEKSI ATRIBUT WEIGHTS BY INFORMATION GAIN  
DAN SELECT BY WEIGHTS PADA ALGORITMA NAÏVE BAYES UNTUK  
PREDIKSI KOLEKTIBILITAS PEMBIAYAAN USAHA KECIL DAN  
MENENGAH**

*Agus Alim Muin S.Kom, M.Kom*  
(*alim.blues@gmail.com*)

**ABSTRAK**

*Kredit adalah pinjaman uang dengan pembayaran tidak tunai dimana angsuran pembayarannya wajib dilakukan oleh peminjam kepada bank atau badan lain sesuai perjanjian kedua belah pihak dan dalam jangka waktu yang sudah di setujui bersama khususnya dalam hal ini adalah melirik kepada sebuah badan usaha kecil dan menengah.*

*Usaha kecil dan menengah (UKM) adalah suatu badan usaha yang dibuat oleh organisasi atau bisa juga oleh perorangan agar dapat menciptakan suatu bidang usaha dan membuat lapangan kerja sehingga dapat menghasilkan finansial serta diharapkan dapat membantu membangun perekonomian Indonesia pada umumnya, Usaha kecil dan menengah (UKM) ingin agar dapat cepat mendapatkan dana kredit dari bank atau proposal usahanya ingin cepat di setujui oleh bank, sehingga arus timbal balik dari aspek ini menyebabkan bank membuat suatu kebijakan kelayakan kredit sesuai kolektibilitas bagi usaha kecil dan menengah (UKM) agar nantinya bank dapat mengetahui apakah usaha kecil dan menengah masuk ke dalam salah satu kategori di kolektibilitas.*

*Dari permasalahan tersebut digunakan sebuah metode klasifikasi sekaligus memprediksi kolektibilitas kredit usaha kecil dan menengah (UKM) dengan model algoritma Naive bayes berbasis weight information gain dan select by weight Setelah dilakukan pengujian dengan model Algoritma naive bayes berbasis weight information gain dan select by weights menghasilkan akurasi sebesar 84.64%.*

**Kata Kunci** : : *Naive bayes, Prediksi, kolektibilitas, select by weights, weight information gain.*

**PENDAHULUAN**

Kredit adalah pinjaman uang dengan pembayaran tidak tunai dimana angsuran pembayarannya wajib dilakukan oleh peminjam kepada bank atau badan lain sesuai perjanjian kedua belah pihak dan dalam jangka waktu yang sudah di setujui bersama khususnya dalam hal ini

adalah melirik kepada sebuah badan usaha kecil dan menengah.

Menganalisa kolektibilitas kredit merupakan hal yang penting dalam lingkup resiko keuangan. karena dapat mempengaruhi perkembangan bank itu sendiri, oleh karena itu perlu dilakukan analisa.

Data yang besar dan banyaknya variable tentunya membutuhkan alat yang efektif dan efisien untuk

melakukan analisa kredit dan menilai debitur yang mempunyai resiko gagal bayar dan yang tidak beresiko.

**RUMUSAN MASALAH**

Berdasarkan permasalahan diatas maka rumusan masalah penelitian ini adalah penerapan algoritma naïve bayes berbasis *weights by information gain* dan *select by weights* untuk prediksi kolektibilitas usaha kecil dan menengah. Sehingga “Seberapa besar peningkatan akurasi metode Naive Bayes dengan menerapkan seleksi atribut berbasis *weights informaion gain* dan *select by weights* untuk prediksi kolektibilitas usaha kecil dan menengah?”

**TUJUAN PENELITIAN**

Berdasarkan latar belakang dan rumusan masalah diatas, maka penelitian ini bertujuan untuk menerapkan metode seleksi atribut *Weights information gain* dan *select by weights* untuk meningkatkan akurasi dari algoritma klasifikasi *Naive Bayes* dalam memprediksi kolektibilitas usaha kecil dan menengah sehingga didapatkan tingkat akurasi terbaik.

**TARGET LUARAN DAN MANFAAT PENELITIAN**

Target luaran dalam pembuatan laporan penelitian ini adalah untuk mempelajari dan untuk memfasilitasi pemahaman tentang model algoritma *naive bayes berbasis weights information gain* dan *select by weight*.

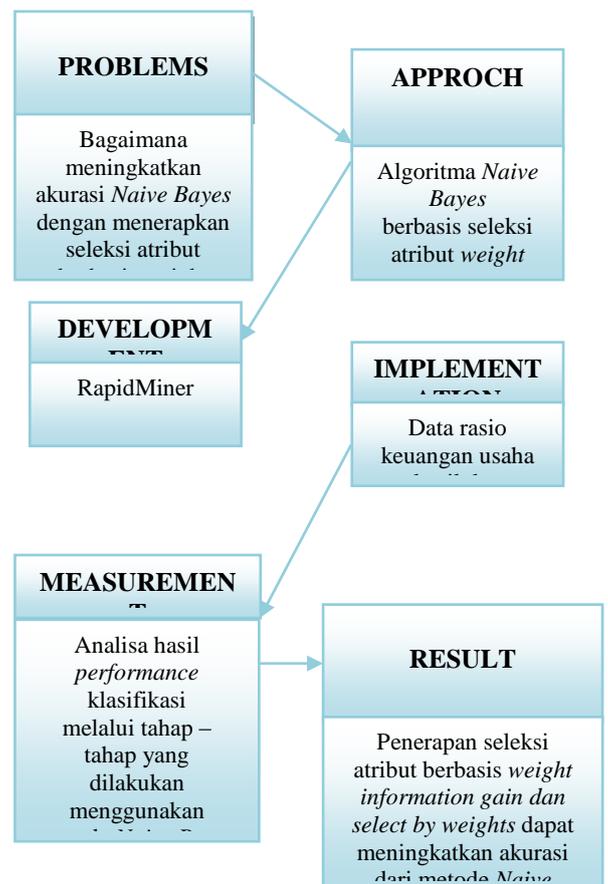
**METODE PENELITIAN**

Pada dasarnya Penerapan seleksi atribut *weights by information Gain* dan *select by weights* pada algoritma *naive bayes* untuk prediksi kolektibilitas pembiayaan usaha kecil dan menengah

terdiri dari tiga bagian.

- a. Pengolahan data awal
- b. Metode yang diusulkan
- c. Eksperimen dan pengujian metode

Pengukuran kinerja dilakukan dengan menghitung akurasi dari prediksi. Semakin besar nilai akurasi dari masing-masing parameter ini menyatakan semakin dekat nilai prediksi dengan nilai sebenarnya. Dengan demikian dapat diketahui algoritma yang lebih akurat.



Gambar 1 Proses penerapan seleksi atribut berbasis *weight information gain* dan *select by wiegths*

**1. Pengumpulan Data Awal**

Data yang didapatkan dari instansi terkait masih berupa data yang terdiri dari Data keuangan dan ratio keuangan usaha nasabah, sehingga harus direkapitulasi terlebih dahulu. Rekapitulasi tersebut dilakukan dengan memperhatikan segmentasi nasabah. Data hasil proses ini adalah data dengan 24 atribut.

**2. Metode yang Diusulkan**

Metode yang diusulkan adalah algoritma *naive bayes* dengan seleksi atribut *weights by information gain* dan *select by weights* untuk memprediksi data ratio keuangan nasabah untuk pengambilan keputusan. Algoritma *naive bayes* dengan seleksi data akan implementasikan dengan menggunakan RapidMiner 5.1.001.

**ANALISA HASIL DAN PEMBAHASAN**

**1. Hasil Penelitian**

Komputer yang digunakan untuk menjalankan penelitian ini adalah laptop, dengan prosesor AMD dual core E1 1.5 GHz dan RAM 2 GB, dengan system operasi Microsoft windows 8. Pada tahap pertama pengujian dilakukan dengan menggunakan algoritma *naive bayes* dengan dengan validasi model klasifikasi dilakukan terhadap data *testing* dengan teknik *10-folds X-validation*.

Tabel 1 Hasil akurasi *naive bayes*

Number of Validation	Akurasi
2	84.60%
3	84.62%
4	84.61%
5	84.63%
6	84.61%
7	84.63%
8	84.63%

9	84.67%
10	84.62%

hasil tersebut belum maksimal karena pada kenyataanya ada artibut yang tidak mempunyai pengaruh dalam menentukan hasil akurasi. Oleh sebab itu maka untuk membuat model keputusan yang baik di tambahkan lagi seleksi atribut *weights by information gain* dan *select by weights*. *10-folds X- validation* hasil penambahan atribut:

Number of Validation	Akurasi
2	85.67%
3	85.67%
4	85.85%
5	85.86%
6	85.87%
7	85.86%
8	85.85%
9	85.86%
10	85.87%

**2. Implementasi Sistem**

Implementasi sistem program ini mencakup spesifikasi kebutuhan perangkat keras (*hardware*) dan spesifikasi perangkat lunak (*software*).

**3. Spesifikasi Perangkat Keras dan Perangkat Lunak**

Program ini direkomendasikan untuk dijalankan dengan menggunakan perangkat keras (*hardware*) yang mempunyai spesifikasi berikut:

1. Prosesor Minimal Intel Pentium atau AMD.
2. Memory minimal 1 GB.
3. Harddisk minimal 320 GB.
4. VGA *card* 64 MB. (Optional)
5. Monitor dengan resolusi 1024 × 768 *pixel*.
6. *Keyboard* dan *Mouse*.

Adapun perangkat lunak (*software*) yang digunakan untuk menjalankan aplikasi ini adalah lingkungan sistem operasi minimal *MS-WindowXP* atau *Windows 7* dan *Windows 8*.

#### 4. Pengujian Program

Pada tahap pertama pengujian dilakukan dengan menggunakan algoritma *naive bayes* dengan dengan validasi model klasifikasi dilakukan terhadap data *testing* dengan teknik *10-folds X-validation*

accuracy : 84.63% +/-7.59% (mikro: 84.62%)							
	true 1	true 2	true 3	true 4	true 5	class precision	
pred 1	341	44	8	4	6	84.62%	
pred 2	0	0	0	0	0	0	
pred 3	0	0	0	0	0	0	
pred 4	0	0	0	0	0	0	
pred 5	0	0	0	0	0	0	
class recall	100%	0.00%	0.00%	0.00%	0.00%		

Gambar 1. Hasil Pengujian data *naive bayes* X- validasi=10

Selain pengujian data *naive bayes* dengan X-validasi 10 ada hasil selanjutnya yaitu X-validasi 9 dimana keduanya diuji dengan menggunakan algoritma *naive bayes* yang hasilnya

accuracy : 84.67% +/-10.11% (mikro: 84.62%)							
	true 1	true 2	true 3	true 4	true 5	class precision	
pred 1	341	44	8	4	6	84.62%	
pred 2	0	0	0	0	0	0	
pred 3	0	0	0	0	0	0	
pred 4	0	0	0	0	0	0	
pred 5	0	0	0	0	0	0	
class recall	100%	0.00%	0.00%	0.00%	0.00%		

Gambar 2. Hasil Pengujian data *naive bayes* X- validasi=9

Dari kedua X-validasi tersebut hasil tersebut belum maksimal karena pada kenyataannya ada artibut yang tidak mempunyai pengaruh dalam menentukan hasil akurasi maka untuk membuat model keputusan yang baik di tambahkan lagi seleksi atribut *weights by information gain* dan *select by weights* dengan X validasi 10 dan 9

accuracy : 85.87% +/-4.21% (mikro: 85.86%)							
	true 1	true 2	true 3	true 4	true 5	class precision	
pred 1	338	44	0	4	6	86.22%	
pred 2	0	0	0	0	0	0	
pred 3	3	0	8	0	0	72.73%	
pred 4	0	0	0	0	0	0	
pred 5	0	0	0	0	0	0	
class recall	99%	0.00%	100.00%	0.00%	0.00%		

Gambar 4. Hasil Pengujian data *naive bayes* dengan seleksi atribut X- validasi= 10

Dari penambahan seleksi atribut *weights by information gain* dan *select by weights* X validasi 10 mengalami penambahan nilai akurasi dan X validasi 9 juga mengalami pembahan nilai akurasi

accuracy : 85.86% +/-4.03% (mikro: 85.86%)							
	true 1	true 2	true 3	true 4	true 5	class precision	
pred 1	338	44	0	4	6	86.22%	
pred 2	0	0	0	0	0	0	
pred 3	3	0	8	0	0	72.73%	
pred 4	0	0	0	0	0	0	
pred 5	0	0	0	0	0	0	
class recall	99%	0.00%	100.00%	0.00%	0.00%		

Gambar 4. Hasil Pengujian data *naive bayes* dengan seleksi atribut X- validasi= 9

## PENUTUP

### 1.KESIMPULAN

Dari hasil penelitian yang dilakukan dari tahap awal hingga pengujian, dan hasil perbandingan dapat disimpulkan bahwa, Penerapan seleksi atribut dapat meningkatkan akurasi *Naive bayes* di x validasi 9 menjadi lebih baik dari 84.67 % menjadi 85.86 %. Meskipun kenaikan yang dihasilkan tidak terlalu besar yaitu hanya sebesar 1,19 % namun secara umum hasil penerapan seleksi atribut menggunakan *Naive bayes* berbasis *weights by information gain* lebih baik daripada *Naive bayes* tanpa menggunakan seleksi atribut.

### 2. Saran

Penelitian hanya uji coba untuk menilai tingkat akurasi penerapan *Naive bayes* berbasis *weights information Gain* dan *select by weights* pada prediksi

kolektibilitas Usaha kecil dan menengah (UKM). Agar penelitian ini dapat memberikan kontribusi yang lebih besar, perlu untuk pengembangan dengan cara mencoba dengan Algoritma lain dan seleksi atribut yang laen, sehingga bisa mendapatkan hasil yang lebih baik untuk penyelesaian masalah dalam dalam memprediksi usaha kecil dan menengah.

#### DAFTAR PUSTAKA

- [1] M. Dunham, Data Mining Introductory and Advanced Topics, New Jersey: Prentice Hall, 2003
- [2] Larose, Data Mining Methods And Models, Canada: John Wiley & Sons, Inc, 2006.
- [3] C. Sammut dan G. Web, Encyclopedia of machine learning, New York: Springer, 2011.
- [4] Zhang dan Wang, Application of Bayesian Method to Spam SMS Filtering, International Conference
- [5] J. Zurada dan K. N. Kunene, "Comparisons of the Performance of Computational Intelligence Methods for Loan Granting Decisions," dalam *Proceedings of the 44th Hawaii International Conference on System Sciences 2011*,
- [6] M. Mahmoud dan A. A. Najla Algadi, "Expert System for Banking Credit Decision," dalam *International*
- [7] M. Mahmoud dan A. A. Najla Algadi, "Expert System for Banking Credit Decision," dalam *International* Mitchell. (1997). Machine Learning. McGraw Hill.
- [8] J. J. R. Martha C Polson, Foundation of Intelligent Tutoring Systems, New Jersey: Lawrence Erlbaum Associates
- [9] a. S. R. Jim Ong, "Intelligent Tutoring Systems : The What and The How".
- [10] A. T. C. K. R. K. Ryan Shaun Baker, "Detecting Student Misuse of Intelligent Tutoring Systems" ..