

PREDIKSI CHURN DENGAN ALGORITMA SELF ORGANIZING MAP'S KOHONEN (SOM) DAN BACKPROPAGATION (BP)

Ziadatul Maskana Ulfa¹, Agung Toto Wibowo², Zk. Abdurahman Baizal³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Churn prediction merupakan salah satu jenis task data mining, yaitu klasifikasi yang bertujuan untuk memprediksi pelanggan yang berpotensi untuk churn. Dalam tugas akhir ini digunakan dua metode Jaringan saraf tiruan untuk melakukan prediksi pelanggan yang churn. Algoritma pertama adalah algoritma Backpropagation, dimana algoritma ini mempunyai keakuratan prediksi yang tinggi. Algoritma kedua adalah algoritma Self Organizing Maps Kohonen, dimana algoritma ini merupakan algoritma yang bagus digunakan untuk klusterisasi data yang dapat dimanfaatkan untuk mengelompokkan data berdasarkan pola-pola data yang dipelajari. Berdasarkan fungsi dari masing-masing algoritma tersebut, pada tugas akhir ini akan algoritma SOM-BP dimana algoritma ini merupakan kombinasi dari kedua algoritma diatas. Data yang digunakan pada tugas akhir ini adalah sample data Tournament. Dalam Tugas Akhir ini akurasi yang dihasilkan diukur dengan tiga parameter yaitu lift curve, top decile dan f-measure. Untuk data undersampling, pengukuran lift curve terbaik 10% customer SOM bisa menebak 59% , pengukuran top decile SOM 10% customer sebesar 1.3 dan pada pengukuran f-measure terbaik yaitu SOM-BP dengan nilai 0.3991.

Kata Kunci : Jaringan saraf tiruan, Algoritma Backpropagation Network, Algoritma Self Organizing Maps Kohonen, Prediksi Churn, Lift curve, Top Decile, f-measure.

Abstract

Churn prediction is one type of data mining tasks, namely classification which aims to predict the potential for customer churn. In this thesis used two methods of artificial neural networks to predict customer churn. The first algorithm is the Backpropagation algorithm, which algorithm has a high prediction accuracy. The second algorithm is an algorithm Kohonen Self Organizing Maps, where this algorithm is a good algorithm is used to klusterisasi data that can be used to classify data based on data patterns are studied. Based on the function of each of these algorithms, in this final SOM-BP algorithm where the algorithm is a combination of the two algorithms above. Data used in this thesis is the sample data Tournament. In this final accuracy of the resulting measured by three parameters namely the lift curve, the top decile and f-measure. For data undersampling, measurement of the lift curve is the best 10% customer can guess 59% SOM, measuring 10% of top decile SOM customer of 1.3 and the best measurement of f-measure of SOM-BP with a value of 0.3991.

Keywords : Artificial Neural Network, Backpropagation Network Algorithm, Self Organizing Maps Kohonen Algorithm, Churn Prediction, , Lift curve, Top Decile, f-measure.

1. Pendahuluan

1.1 Latar belakang

Asset perusahaan yang paling bernilai adalah pelanggan setia yang mereka miliki. Oleh karena itu, perusahaan akan terus berusaha untuk meningkatkan kesetiaan dan kepuasan pelanggannya. Namun, dalam suatu bisnis, meningkatkan kesetiaan dan kepuasan pelanggan adalah hal yang sulit. Dengan semakin tingginya kebutuhan masyarakat, maka semakin banyak perusahaan baru yang muncul dengan keberagaman layanan yang ditawarkan. Ketatnya persaingan bisnis antara perusahaan, membuat pelanggan memiliki banyak pilihan dan dapat dengan mudah pelanggan melakukan *churn*. *Churn* merupakan suatu ukuran jumlah orang atau benda yang masuk atau keluar dari suatu kelompok dalam jangka waktu tertentu. Pengertian *churn* disini yaitu ketika semua jasa perusahaan yang digunakan seorang pelanggan diputus oleh pelanggan ataupun oleh perusahaan. Pada konsep *churn* sendiri dikatakan, “lebih murah cost bagi perusahaan untuk mempertahankan *customer* mereka daripada mengupayakan menarik *customer* yang baru.”[9]. Oleh karena itu, perlu adanya suatu model prediksi yang akurat sehingga dapat memprediksi pelanggan yang akan *churn*. Prediksi *churn* ini akan membantu suatu perusahaan agar tidak mengalami kerugian yang lebih besar.

Prediksi *churn* adalah salah satu aplikasi task *data mining* yang bertujuan untuk memprediksi pelanggan yang berpotensi untuk *churn*. Hal ini dilakukan agar suatu perusahaan dapat memiliki analisis bisnis dan identifikasi karakteristik pelanggan yang tepat dalam meningkatkan kompetisi demi keuntungan perusahaan, menguasai *market share* pada bidang pemasaran dan perusahaan dapat melakukan penanggulangan yang tepat kepada karakteristik pelanggan tersebut. Namun, karakteristik dari data *churn* adalah tingkat *imbalance* yang besar, karena pelanggan yang mengalami *churn* jauh lebih sedikit dibandingkan dengan pelanggan yang loyal. Ini mengakibatkan sulitnya membuat pemodelan terhadap data *churn* (Cardell,2003). Oleh karena itu perlu dilakukan *balance* data dengan metode *sampling*.

Pada metode ini akan digunakan algoritma-algoritma jaringan saraf tiruan untuk memprediksi pelanggan yang *churn* berdasarkan data *history* pelanggan yang sudah ada. Salah satu algoritma pembelajaran yang dapat diaplikasikan dalam bidang peramalan / *forecasting* adalah *Backpropagation* [2]. Sebagaimana halnya dengan suatu algoritma yang dapat dimodifikasi ataupun dikombinasi dengan algoritma lainnya, algoritma *Backpropagation* juga dapat dimodifikasi ataupun dikombinasi dengan algoritma lainnya. Algoritma pembelajaran lainnya untuk memprediksi pelanggan *churn* yang akan diuji pada penelitian ini adalah algoritma *Self Organizing Maps Kohonen (SOM)*. Berbeda dengan algoritma *Backpropagation* yang melakukan pembelajaran dengan pengawasan (*supervised*), algoritma SOM melakukan pembelajaran tanpa pengawasan (*unsupervised*). Algoritma SOM merupakan salah satu teknik *clustering* yang dapat mengelompokkan data kedalam pola-pola tertentu secara *unsupervised* [2]. Untuk mendapatkan hasil yang lebih baik dalam proses prediksi berdasarkan sifat-sifat algoritma SOM dan BP tersebut, maka pada penelitian ini akan menggunakan kombinasi algoritma *Backpropagation* dengan algoritma *Self Organizing Maps Kohonen*, yang dikenal dengan metode *Self Organizing Maps Kohonen*-

Backpropagation Jaringan saraf tiruan (SOM-BP). Algoritma *SOM* sebagai lapisan awal untuk memproses input, kemudian algoritma *Backpropagation* akan memproses sinyal input hasil dari proses pembelajaran algoritma *SOM* sebagai *factor pembimbing* untuk memprediksi *Churn* di masa yang akan datang. Dengan kata lain algoritma *SOM* digunakan untuk menyaring data ke dalam pola-pola dan menyederhanakan proses pembelajaran algoritma *Backpropagation* untuk mendapatkan keakuratan yang lebih baik dalam proses prediksi dibandingkan dengan keakuratan algoritma *SOM* dan *BP* itu sendiri.

Hasil akhir dari tugas akhir ini yaitu dapat mengetahui bagaimana keakuratan prediksi data *Churn* dengan menggunakan metode *Self-organizing map's kohonen (SOM)* dan *Backpropagation (BP) Jaringan saraf tiruan* dibandingkan dengan algoritma *Self-organizing map's kohonen (SOM)* itu sendiri dan algoritma *Backpropagation (BP)* itu sendiri yang penghitungan akurasi modelnya dinyatakan dalam bentuk *Top decile*, *Lift Curve* dan *f-measure*. Jika dilihat dari sudut pandang operator, akurasi difokuskan pada 10% *riskiest segment* yaitu pengukuran *lift curve* dan *top decile*. Pertimbangan dalam memilih nilai 10% adalah karena kelompok yang meliputi 10% *customer* dengan tingkat resiko tertinggi merupakan segmentasi ideal bagi perusahaan dalam menerapkan strategi pemasaran untuk mencegah terjadinya *churn* [14]. Oleh karena itu, dalam hal ini baik yang dimaksudkan dalam hipotesa awal adalah akurasi terbaik berdasarkan pengukuran *top decile*.

1.2 Perumusan Masalah

Dengan mengacu pada masalah latar belakang diatas, maka permasalahan yang akan dibahas dan diteliti adalah:

1. Bagaimana merancang sistem prediksi *churn* dengan menggunakan *Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan*.
2. Bagaimana akurasi model prediksi *churn* yang didapat dari proses klasifikasi data hasil penerapan metode *Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan* yang dinyatakan dalam bentuk *top decile*, *lift curve* dan *f-measure*.

1.3 Tujuan

Berdasarkan rumusan masalah diatas, maka tujuan dari tugas akhir ini adalah:

1. Menghasilkan suatu aplikasi yang mengimplementasikan metode *Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan* untuk prediksi *churn*.
2. Membandingkan tingkat akurasi metode *Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan* dengan *Self-organizing map's kohonen (SOM)* dan *Backpropagation (BP)* yaitu melalui perhitungan tingkat keberhasilan dalam parameter *top decile*, *lift curve* dan *f-measure*.

Batasan masalah dalam tugas akhir ini adalah:

1. Data yang digunakan adalah *sample data Tournament 2003*.
2. *Preprocessing* data dilakukan dengan menggunakan *tool Clementine 12.0*.

3. Aplikasi ini merupakan aplikasi yang berdiri sendiri, tidak menempel pada *tool Clementine 12.0*.
4. Proses *Balancing* dilakukan dengan menggunakan *tool Clementine 12.0*.

1.4 Metodologi Penyelesaian Masalah

Metode yang digunakan dalam penyelesaian tugas akhir ini dengan langkah kerja sebagai berikut:

1. Studi Literatur :
 - a. Mencari referensi dan sumber-sumber lain dari beberapa buku, jurnal, artikel yang membahas tentang *data mining, churn prediction, imbalance class, matlab programming, clementine 12.0, algoritma Self Organizing Maps Kohonen (SOM), algoritma Backpropagation, Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan* dan pengukuran evaluasi untuk prediksi.
 - b. Pendalaman materi, mempelajari dan memahami materi yang berhubungan dengan tugas akhir.
2. Pencarian dan pemahaman data pelanggan *churn* agar sebagai awal pengetahuan untuk pemodelan prediksi *churn*.
3. Melakukan analisa permasalahan yang akan diselesaikan dengan metode *Self Organizing Maps Kohonen (SOM), algoritma Backpropagation* dan *Self-organizing map's kohonen - Backpropagation (SOM-BP) Jaringan saraf tiruan*.
4. Melakukan analisa dan perancangan perangkat lunak.
5. Melakukan Implementasi perangkat lunak dengan *Matlab R2009a* dan melakukan *preprocessing* data dengan *tool Clementine 12.0*.
6. Melakukan testing terhadap sistem aplikasi yang telah dibangun dengan data yang telah dikumpulkan dan di *preprocessing* sebelumnya.
7. Pengambilan kesimpulan dan penyusunan laporan tugas akhir.

5. Penutup

5.1 Kesimpulan

1. SOM-BP memberikan hasil prediksi yang baik untuk pengujian data hasil *balancing* dengan menggunakan metode undersampling.
2. Akurasi SOM-BP dengan metode undersampling mendapatkan akurasi untuk keseluruhan customer lebih tinggi dibandingkan dengan metode lain, namun tidak untuk 10% riskiest segment.
3. Algoritma SOM dan Algoritma BP memberikan hasil yang baik di pengujian yang melibatkan *riskiest segment*.
4. Evaluasi akurasi terbaik yaitu pada evaluasi *Top decile* dimana nilai *Top decile* tertinggi pada algoritma SOM.

5.2 Saran

1. Lebih teliti lagi dalam melakukan *preprocessing data*, karena keakuratan hasil prediksi bukan hanya ditentukan dari metode yang digunakan, tetapi juga dipengaruhi oleh pengolahan *preprocessing data*.
2. Sistem *Churn Prediction* dapat dikembangkan lagi dengan parameter keakuratan waktu, misalnya dapat menebak banyaknya pelanggan yang akan *churn* dalam periode tertentu.
3. Dikembangkan dan dimodifikasi lagi sistem pembelajaran *SOM-BP* untuk mendapatkan hasil yang lebih akurat dibandingkan SOM dan BP itu sendiri dalam sudut pandang operator.

Daftar Pustaka

- [1] Siang.Jong-jek,"*Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan Matlab*". Andi Yogyakarta. 2005.
- [2] Jiawei Han and Micheline Kamber, "*Data Mining : Concepts and Techniques*", Intelligent Database Systems Research Lab, School of Computing Science, Simon Fraser University.
- [3] Hand.David, Mannila.Hellki, Smyth.Padmraic, "*Principles of data mining*", Massachusets Institut of Technology, 2001.
- [4] Lubis .Chairisni, Suttedjo. Eddy, Setiadi .Bowo, "*Prediksi Harga Saham dengan Menggunakan Algoritma Hybrid Neural Network*", Universitas Tarumanagara, Yogyakarta,2005.
- [5] Suyanto, "*Soft Computing*", Informatika Bandung, 2008
- [6] Suyanto, "*Artificial Intellegence*", Informatika Bandung , 2007
- [7] Tan.Pang-Ning, Steinbach. Michael, Kumar.Vipin, "*Introduction to Data Mining*", Peorson Education: America, 2006.
- [8] Kusnawi, "*Pengantar Solusi Data Mining*", Seminar Teknologi Yogyakarta: Yogyakarta, 2007.
- [9] John Hadden, Ashutosh Tiwari, Rajkumar Roy, and Dymitr Ruta,. *Churn Prediction: Does Technology Matter?*, ISSN 2006.
- [10] Haykin, Simon. "*Neural Network, A Comprehensive Foundation*". Macmillan Publishing Company, USA. 1994.
- [11] "*Network Design & Training*", CogNova Tekhnologi, ESCOM, 2010
- [12] Kennedy, "*Solving data mining problems*".
- [13] Baizal. ZK. Abdurahman, Bijaksana. Moch. Arif, Sagita Sastrawan.Angelina," *Analisis Pengeruh Metode Oversampling dalam Churn Prediction Untuk Perusahaan Telekomunikasi*". Seminar Nasional Aplikasi Teknologi Informasi, Yogyakarta , 2009.
- [14] Lemmend, Aurelie and Christopher Croux. *Bagging and Boosting Classification Trees to Predict Churn*. Departement of Appliend Economics,K.U. Leuvan, Belgium.

- [15] “*Neural Network*” . <http://www.statsoft.com/textbook/neural-networks/#preandpost> diakses pada tanggal 16 Desember 2010.

