

IMPLEMENTASI DAN ANALISA BEBERAPA METODE KLASIFIKASI UNTUK MENENTUKAN ATURAN PADA LOKASI RAWAN KECELAKAAN IMPLEMENTATION AND ANALYSIS SOME CLASSIFICATION METHOD TO DETERMINE RULE IN TRAFFIC ACCIDENT LOCATION

Iwan Budy Antono^{1, -2}

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Era komputerisasi telah membuka wawasan dan paradigma baru dalam proses pengambilan keputusan dan penyebaran informasi. Di Indonesia, pada saat ini, Sistem Informasi Geografis (SIG) telah dikenal secara luas sebagai alat bantu untuk pengambilan keputusan. Dengan adanya SIG, pengguna dapat memvisualisasikan dan menganalisa suatu area studi berdasarkan lokasi - lokasi unsur - unsur geografi tertentu, misalnya dalam menentukan relasi atau hubungan (relationship), pola (pattern), kecenderungan (trend) pada basisdata lokasi rawan kecelakaan lalu - lintas.

Klasifikasi adalah salah satu teknik dari data mining yang sering digunakan untuk menentukan aturan (rule) dan pola (pattern) yang muncul dari basisdata tertentu. Teknik yang umum digunakan dalam suatu metode Klasifikasi adalah suatu pendekatan sistematis untuk menentukan model klasifikasi dari sekumpulan data.

Tujuan dari tugas akhir ini adalah untuk mengimplementasikan dan menganalisa beberapa metode Klasifikasi, yaitu : Decision Tree, Rule Based Classifier, dan Naive Bayessian yang akan diterapkan pada basisdata Sistem Informasi Geografis untuk meningkatkan hasil prediksi dalam hal penentuan lokasi yang rawan kecelakaan sebagai dasar dalam pengambilan keputusan selanjutnya.

Kata Kunci : Analisa Klasifikasi, Decision Tree, Rule Based Classifier, Naive Bayessian, Sistem Informasi Geografis (SIG)

Abstract

Computerize era have opened knowledge and new paradigm in course of decision making and information spreading. In Indonesia, at this time, Geographical Information System (GIS) have been recognized widely as a means of assist for decision making. With existence GIS, user can be visualized and analyse a study area of pursuant to certain location element of geographi, for example in determining relationship, pattern, tendency at database of troubled traffic accident location.

Classification is one part of data mining which often be used to determine rule and pattern what emerge from database of Geographical Information System. Technique in Classification method is a systematic approach to determine classification model from a group of data.

Intention of this final assignment is to implemented and analyse some classification method, that is : Decision Tree, Rule Based Classifier, and Naive Bayessian to be applied at database of Geographical Information System to increase prediction result in the case of location determinated which the troubled traffic accident as basic in decission making furthermore.

Keywords : Classification Analysis, Decision Tree, Rule Based Classifier, Naive Bayessian, Geographic Information System (GIS)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di Indonesia, pada saat ini, masalah penentuan lokasi rawan kecelakaan lalu lintas adalah sangat penting. Seorang pakar jalan raya atau perencana kota dapat menampilkan segmen – segmen jalan mana saja yang sangat rawan kecelakaan berdasarkan informasi kondisi segmen jalan tersebut. Kemudian memutuskan segmen jalan mana saja yang memerlukan perbaikan. Untuk memperoleh tingkat akurasi yang tinggi maka akan sering dilakukan pemantauan secara on-line melalui kamera atau citra satelit, selain memerlukan biaya yang mahal, faktor buruk lainnya adalah tingkat akurasi yang sangat rendah dalam memprediksi penentuan lokasi yang rawan kecelakaan di masa yang akan datang.

Data mining adalah proses penyaringan data dimana sebelumnya tidak diketahui terdapatnya informasi yang potensial. Data mining bertujuan untuk menemukan pola dan aturan dalam basis data yang berukuran besar agar dapat digunakan untuk mendukung pengambilan keputusan. Salah satu teknik yang populer dalam data mining adalah Klasifikasi, tujuan dari teknik ini adalah mengelompokkan suatu objek ke dalam kelas tertentu berdasarkan pola kelas tertentu. Decision tree, Rule Based, dan Naïve Bayesian adalah beberapa metode Klasifikasi yang paling sering dipakai karena mudah untuk diinterpretasi dalam kehidupan nyata.

Dengan diterapkannya beberapa metode Klasifikasi seperti Decision Tree, Rule Based, dan Naïve Bayesian pada proses penentuan lokasi rawan kecelakaan tersebut, keuntungan yang didapat selain biayanya yang relatif murah dan waktu proses yang cepat untuk ukuran data yang berukuran besar, metode ini juga memiliki tingkat akurasi yang relatif tinggi dibandingkan dengan metode klasifikasi yang lain dalam hal memprediksi lokasi rawan kecelakaan di waktu yang akan datang.

1.2 Perumusan Masalah

Dalam penyusunan tugas akhir ini dapat dirumuskan beberapa permasalahan pokok antara lain :

- a. Bagaimana mengimplementasikan beberapa metode Klasifikasi, yaitu Decision Tree, Rule Based, dan Naive Bayesian pada Basis Data lokasi rawan kecelakaan lalu – lintas dalam bentuk Perangkat Lunak.
- b. Bagaimana menganalisa dataset lokasi rawan kecelakaan dari beberapa metode diatas untuk menentukan aturan (rule) pada penentuan lokasi rawan kecelakaan lalu – lintas.
- c. Dari hasil analisa sebelumnya, untuk menguji keabsahan rule dari masing – masing metode tersebut, maka akan dilakukan uji coba untuk mengetahui tingkat hasil akurasi dari rule tersebut.

1.3 Tujuan Penelitian

Berdasarkan pada masalah yang telah didefinisikan di atas, maka tujuan dari Tugas Akhir ini adalah mengimplementasikan beberapa metode Klasifikasi, yaitu : Decision Tree, Rule Based, dan Naive Bayesian dalam proses menentukan serta menganalisa aturan (rule) dari data yang terdapat pada basisdata Sistem Informasi Geografis sebagai dasar pengambilan keputusan, Serta untuk mengetahui metode mana yang efektif digunakan dalam hal mengevaluasi dan memprediksi lokasi – lokasi mana saja yang rawan kecelakaan lalu lintas di masa yang akan datang dengan tingkat akurasi ketelitian yang tinggi serta waktu proses yang cepat dan simplisitas rule yang sangat simpel.

1.4 Batasan Masalah

Adapun batasan masalah dalam penyusunan tugas akhir ini antara lain :

- a. Tugas akhir ini hanya difokuskan pada proses pengklasifikasian dalam menentukan rule – rule mining.

- b. Masing – masing metode yang akan digunakan dalam tugas akhir ini akan diinterpretasikan oleh sebuah algoritma saja.
- c. Data yang akan digunakan berasal dari Basis Data lokasi rawan kecelakaan yang telah didiskritisasi.
- d. Dalam tugas akhir ini tidak memperhitungkan tingkat validitas data yang diperoleh dengan kenyataan di lapangan.
- e. Karena Tugas akhir ini hanya fokus pada Analisis Klasifikasi, maka proses *Pre - Processing* tidak ikut diperhitungkan.

1.5 Metodologi Penelitian

Pendekatan sistematis / metodologi yang akan digunakan dalam merealisasikan tujuan dan pemecahan masalah di atas adalah dengan menggunakan langkah – langkah sebagai berikut :

- o Studi Literatur
Mempelajari konsep – konsep Klasifikasi Data Mining khususnya metode Decision Tree, Rule Based, dan Naïve Bayessian serta Sistem Informasi Geografis secara global yang diperoleh dari buku – buku, literatur, jurnal, white paper baik dari dalam maupun luar.
- o Pendalaman materi
Mendalami materi yang akan digunakan seperti Data Mining, metode Analisis Klasifikasi, Metode Decision Tree, Rule Based, dan Naïve Bayessian serta konsep Sistem Informasi Geografis.
- o Perancangan dan implementasi metode
Merancang dan mengimplementasikan hasil rancangan sistem penentuan lokasi rawan kecelakaan menggunakan bahasa pemrograman MS Visual Basic dan DBMS MS SQL Server 2000.
- o Analisa dan Evaluasi
Menganalisa dengan mengukur hasil implementasi melalui perbandingan dengan hasil yang dicapai bila menggunakan metode atau tools lain.

1.6 Sistematika Penulisan

Tugas Akhir ini akan disusun berdasarkan sistematika penulisan sebagai berikut :

- **BAB I : Pendahuluan**
Bab ini membahas antara lain, latar belakang dilakukannya penelitian, perumusan masalah yang akan dianalisa, tujuan yang ingin dicapai melalui penelitian ini, batasan – batasan permasalahan, serta metodologi pemecahan masalah dan sistematika pembahasan.
- **BAB II : Landasan Teori**
Bab ini memuat penjelasan mengenai berbagai dasar teori yang mendukung dan mendasari penulisan tugas akhir ini yaitu mengenai Sistem Informasi Geografi(SIG), Data mining, Klasifikasi.
- **BAB III : Analisa dan Perancangan Sistem**
Bab ini menguraikan spesifikasi terhadap model aplikasi yang akan dibangun. Aplikasi ini dibangun sebagai media pengujian yang bertujuan memahami secara jelas proses yang dilakukan pada aplikasi tersebut sesuai dengan Bab II.
- **BAB IV : Implementasi dan Pengujian Sistem**
Bab ini berisi tentang implementasi aplikasi yang dibangun dan analisa output yang dihasilkan.
- **BAB V : Penutup**
Berisi tentang kesimpulan dari hasil penelitian tugas akhir ini serta saran-saran untuk pengembangan lebih lanjut.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan Pengujian dan Analisa terhadap DataSet Lokasi Rawan Kecelakaan Lalu – Lintas, maka dapat diambil beberapa kesimpulan sebagai berikut :

1. Metode Klasifikasi dapat diterapkan dalam memprediksi Lokasi Rawan Kecelakaan dengan tingkat keakurasian yang cukup tinggi, waktu yang cepat, serta dengan biaya yang relatif murah.
2. Dari ketiga metode klasifikasi yang diimplementasikan dalam tugas akhir ini, metode Decision Tree dan Rule-Based Classifier memiliki grafik waktu eksekusi yang cenderung meningkat secara linier berbanding lurus dengan jumlah record data baik pada saat proses Learning maupun Training dataset. Sedangkan metode Naive Bayessian menunjukkan grafik waktu eksekusi yang relatif konstan.
3. Dalam analisa Simplitas Rule, metode Decision Tree cenderung menghasilkan jumlah rule yang lebih banyak dibandingkan metode Rule-Based Classifier.
4. Dari segi tingkat akurasi, secara umum metode Rule-Based memiliki tingkat akurasi yang lebih tinggi jika dibandingkan dengan metode Decision Tree. Sedangkan metode Naive Bayessian memiliki tingkat keakurasian yang relatif terhadap metode Decision Tree dan Metode Rule-Based Classifier, karena metode Naive Bayessian akan mengkalkulasikan probabilitas semakin tinggi jika jumlah Training Dataset-nya semakin banyak.

5.2 Saran

Untuk pengembangan lebih lanjut dari sistem yang telah dibangun, penerapan beberapa metode dan algoritma klasifikasi diharapkan dapat meningkatkan tingkat akurasi dalam memprediksi dataset kecelakaan lalu – lintas.

Selain itu, untuk mendapatkan hasil akurasi yang tinggi maka Data Set yang harus digunakan haruslah berjumlah besar. Sebagai konsekuensinya, waktu pemrosesan akan semakin lama. Hal ini dapat diatasi dengan menggunakan spesifikasi Hardware yang tinggi.



DAFTAR PUSTAKA

- [01] Anders, Karl – Heinrich, 2001, “*Data Mining for Automated GIS data collection*”, Wichmann Verlag, Heidelberg.
- [02] Buttenfield, Barbara, “*Geospatial Data Mining and Knowledge Discovery*”, A UCGIS white paper on emergent research themes, University of Colorado.
- [03] Demers, M. N, 1997, “*Fundamentals of Geographic Information Systems*”, New York : John Willeys & Sons, Inc.
- [04] Gahegan, Mark, “*Data Mining and Knowledge Discovery in the Geographical Domain*”, Pennsylvania : Department of Geography, The Pennsylvania State University.
- [05] Han, J and M. Kamber, 2001, “*Data Mining : Concept and Techniques*”, San Francisco, CA : Morgan Kaufmann, 550 pp, ISBN 1-55860-489-8
- [06] Heryanti, Diah Noor, 2005, “*Data Mining Classification Dengan Metode Decission Tree Menggunakan Algoritma C.45*”, Bandung : Jurusan Teknik Informatika, Sekolah Tinggi Teknologi Telkom
- [07] Kadir, Abdul, 2002, “*Konsep & Tuntunan Praktis Basisdata*”, Yogyakarta : ANDI Yogyakarta
- [08] Kadir, Abdul, 2002, “*Panduan Praktis Pemrograman Visual Basic 6.0 Tingkat Lanjut*”, Yogyakarta : ANDI Yogyakarta.
- [09] Miller, Harvey J, “*Geographic Data Mining and Knowledge Discovery*”, Salt Lake City : Department of Geography, University of Utah.
- [10] Pang – Ning Tan, Michael Steinbach, and Vipin Kumar, “*Introduction to Data Mining*” (pdf files) <http://www.cse.msu.edu/~cse980/lecture.htm>
- [11] Prahasta, Eddy, 2002, “*Konsep – Konsep dasar Sistem Informasi Geografis*”, Bandung : INFORMATIKA.
- [12] Prambudiono, Iko, 2003, “*Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data*”,

<http://www.ilmukomputer.com/.../iko-datamining.pdf>

- [13] Pressman, Robert. S, 1997, “*Software Engineering : A Practitioner’s Approach*”, fourth edition, New York : McGraw – hill, International Editing Software Engineering Series
- [14] Shaufiah, 2005, “*Klasifikasi Dalam Data Mining Menggunakan Algoritma SLIQ*”, Bandung : Jurusan Teknik Informatika, Sekolah Tinggi Teknologi Telkom.
- [15] Wulandari, Amelia S, 2002, “*Perancangan Sistem Informasi Geografis Penentuan Prioritas Perbaikan Jalan Kota Bandung*”, Bandung : Jurusan Teknik Industri, Sekolah Tinggi Teknologi Telkom.

