PENERAPAN ALGORITMA DECISION TREE C4.5 UNTUK KLASIFIKASI PENENTUAN DAFTAR PRIORITAS PENGEMBANGAN JEMBATAN

Erfan Karyadiputra[1)], Nadiya Hijriana[2)]

[1]Fakultas Teknologi Informasi, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari (Erfan Karyadiputra)
email : erfantsy@gmail.com
[1]Fakultas Teknologi Informasi, Universitas Islam Kalimantan Muhammad Arsyad Al Banjari (Nadiya Hijriana)
email : nadyahijriana@gmail.com

Abstrak

Data Mining merupakan proses mengidentifikasi data dan informasi yang bertujuan untuk membantu mendapatkan informasi yang bermanfaat dan meningkatkan pengetahuan penggunanya. Kelancaran distribusi barang maupun jasa sangat dipengaruhi oleh kondisi infrastruktur jalan maupun jembatan yang menghubungkan beberapa wilayah-wilayah yang secara administrasi merupakan bagian dari Provinsi Kalimantan Selatan. Kondisi jembatan yang rusak tentu dapat menghambat pendistribusian barang maupun jasa sehingga perbaikan maupun pemeliharaan perlu terus dilakukan. Dalam penentuan prioritas pengembangan jembatan agar menghindari adanya tumpang tindih (overlaping) suatu kegiatan dalam satu lokasi, ataupun lokasi yang sama sekali belum tertangani dengan baik, maka dibutuhkan perencanaan pengolahan data dan informasi yang terintegrasi, sehingga data dan informasi perencanaan yang diperoleh tersebut dapat dijamin keakuratan datanya. Banyaknya ruas jembatan yang tersebar di Provinsi Kalimantan Selatan dan terbatasnya anggaran perbaikan ataupun pemeliharaan sehingga instansi terkait perlu menentukan daftar prioritas jembatan yang diprioritaskan untuk disetujui dan dilakukan perbaikan maupun pemeliharaan. Salah satu teknik yang dapat diterapkan untuk mengklasifikasi daftar prioritas pengembangan jembatan yaitu dengan menerapkan metode klasifikasi data mining menggunakan algoritma C4.5. Penerapan metode data mining menggunakan algoritma C4.5 menghasilkan performance akurasi sebesar 82.84% dengan nilai AUC sebesar 0,825 sehingga termasuk dalam kategori good clssification.

***Keywords** : Algoritma C4.5, Data Mining, Klasifikasi*

1. PENDAHULUAN

Kelancaran distribusi barang maupun jasa sangat dipengaruhi oleh kondisi infrastruktur jalan maupun jembatan yang menghubungkan beberapa wilayah administrasi dari Provinsi Kalimantan Selatan. Kondisi jembatan yang rusak tentu dapat menghambat pendistribusian barang maupun jasa dan dapat mempengaruhi pertumbuhan perekonomian suatu daerah sehingga perbaikan maupun pemeliharaan perlu terus dilakukan. Dalam penentuan prioritas pengembangan jalan agar menghindari adanya proyek yang tumpang tindih pada suatu kegiatan dalam satu lokasi, ataupun lokasi yang sama sekali belum tersentuh ataupun belum tertangani dengan baik, maka dibutuhkan perencanaan pengolahan data dan informasi yang terintegrasi, sehingga data dan informasi perencanaan yang diperoleh tersebut dapat dijamin keakuratan datanya.

Banyaknya ruas jembatan yang tersebar di wilayah Provinsi Kalimantan Selatan dan terbatasnya anggaran dana pemerintah untuk melakukan perbaikan ataupun pemeliharaan jembatan sehingga instansi terkait perlu menentukan daftar prioritas pengembangan jembatan yang akan disetujui untuk diprioritaskan dilakukan perbaikan atau pemeliharaan menggunakan beberapa atribut dan kriteria. Salah satu teknik yang dapat diterapkan untuk mengklasifikasi dan mengevaluasi daftar prioritas pengembangan jembatan yaitu dengan menerapkan teknik klasifikasi data mining menggunakan algoritma Decision Tree C4.5 sehingga dapat diketahui *performance* akurasi dalam penentuan daftar prioritas pengembangan jembatan dan mengetahui atribut mana yang paling

berpengaruh. Dengan demikian instansi terkait dapat mengevaluasi dan mengembangkan kebijakan ekonomi serta pembangunan infrastruktur dengan lebih baik.

2. METODE PENELITIAN

a. Pengumpulan Data

Tabel 1. Variabel Data Jembatan Provinsi

Variabel	Keterangan	Kategori	
Y	Status	1	Disetujui
		2	Ditolak
X1	Klasifikasi Jembatan	1	Jembatan Kayu
		2	Box Culvert
		3	Jembatan Girder
		4	Gorong-gorong
		5	Pile Slab
		6	Jembatan Beton
		7	Slab Beton
		8	Jembatan Rangka Baja
		9	Box Kayu
		10	Jembatan Belly
		11	Jembatan log
		12	Jembatan Baja
X2	Jenis Konstruksi	1	Kayu
		2	Beton
		3	Baja
		4	Pile Slab
		5	Bely
		6	Slab Beton
		7	Slab 2
X3	Keterangan Jembatan	1	Aspal Terkelupas
		2	Pagar Rusak
		3	Lantai Diganti Baja
		4	Tidak ada Pagar
		5	Berlobang
		6	Lantai Rusak
		7	Lantai Menurun
		8	Timbunan Tanah
		9	Badan Jalan Menurun
		10	Tanah Longsor
		11	Jembatan Darurat
X4	Kondisi Jembatan	1	Baik
		2	Rusak Ringan
		3	Rusak Berat
		4	Dalam Perbaikan

b. Pengolahan Data Awal

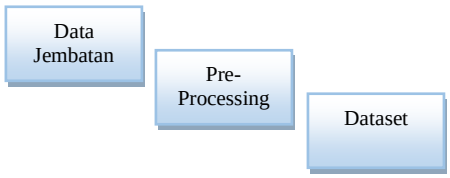
Pada tahap ini, data yang telah dikumpulkan kemudian diolah dengan melakukan *pre-processing* seperti mengidentifikasi dan menginspeksi sejumlah atribut-atribut yang tidak relevan sehingga dapat diproses ketahapan selanjutnya dalam data mining.

c. Eksperimen dan Pengujian Metode

Pada tahap ini dilakukan eksperimen dalam mengimplementasikan pengujian algoritma C4.5. Kemudian selanjutnya membagi dataset menjadi data *training* dan data *testing.* Data *training* digunakan untuk membentuk model klasifikasi sedangkan data *testing* untuk menguji model sehingga diketahui *performance* dari proses klasifikasi tersebut. Tahap selanjutnya melakukan evaluasi dan validasi.

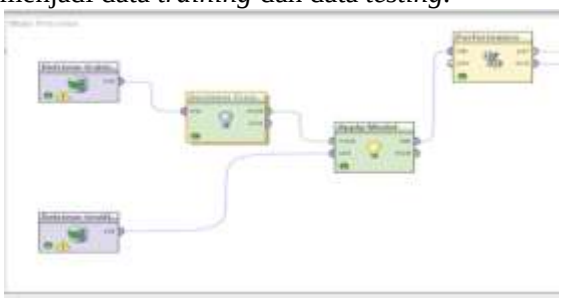
3. HASIL DAN PEMBAHASAN

Sebelum melakukan eksperimen dan pengujian model, data-data yang dikumpulkan terlebih dahulu diolah agar dapat diproses dalam *data mining*.



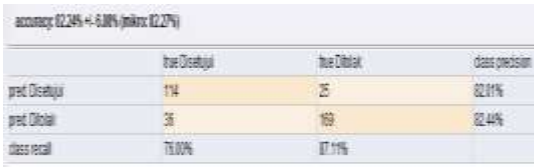
Gambar 1. Pre-Processing

Selanjutnya melakukan implementasi pengujian algoritma C4.5 dan membagi dataset menjadi data *training* dan data *testing*.



Gambar 2. Data training dan testing

Selanjutnya evaluasi dan validasi melalui pengujian *confusion matrix* dan kurva ROC sehingga diketahui akurasi dan modelnya.

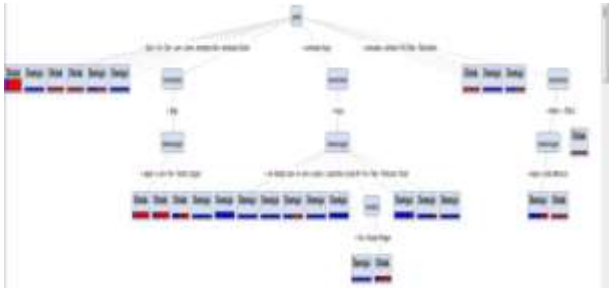


Gambar 3. Hasil Akurasi

Dari hasil pengujian 10-*fold cross-validation*, didapatkan akurasi sebesar 82,84% dengan nilai AUC sebesar 0,825 sehingga termasuk kedalam *good classification* (Gorunescu,2010).

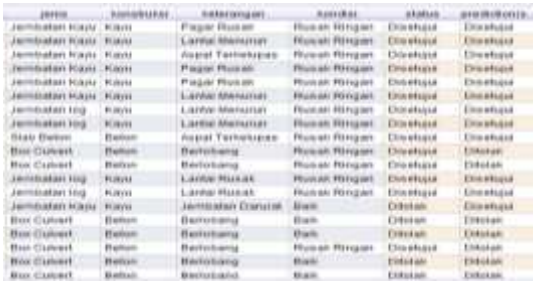


Gambar 4. Kurva ROC



Gambar 5. Model Pohon Keputusan

Sampel data *training* dibandingkan dengan hasil klasifikasi algoritma C4.5 untuk melihat keakuratan hasil klasifikasi dari C4.5.



Gambar 6. Perbandingan Hasil Klasifikasi

Dari hasil perbandingan sampel data yang diuji dengan hasil klasifikasi C4.5 sebagian besar berhasil memprediksi dengan benar dan hanya didapatkan 10 sampel hasil klasifikasi yang tidak sesuai dengan hasil data sampel.

4. KESIMPULAN

Dari hasil penelitian yang dilakukan dari tahap awal pengujian hingga akhir pengujian, maka dapat disimpulkan bahwa penerapan teknik klasifikasi data mining menggunakan algoritma C4.5 bisa diterapkan sebagai salah satu alternatif untuk mengevaluasi daftar prioritas pengembangan jembatan di Provinsi Kalimantan Selatan dengan akurasi sebesar 82,84% dengan nilai AUC sebesar 0,825 sehingga termasuk dalam kategori *good clssification*.

5. REFERENSI

Aprilla, D., Baskoro, D. A., Ambarwati, L., & Wicaksana, I. W. (2013). Belajar Data Mining dengan RapidMiner. Jakarta: academia.edu. Diambil kembali dari http://www.academia.edu

Bellazzi, R., & Zupanb, B. (2008). *Predictive Data Mining In Clinical Medicine: Current Issues And And Guidelines.* International Journal Of Medical Informatics.

Gorunescu, F. (2010). *Data Mining: Concept, Models and Techniques.* Romania: Springer.

Han, J., & Kamber, M. (2007). *Data Mining : Concepts and Techniques* (Second ed.). (M. R. Jim Gray, Penyunt.) San Francisco, United States of America: Morgan Kaufmann Publishers.

Ian H., W., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2 ed.). San Francisco: Morgan Kaufmann Publishers is an imprint of Elsevier.

Larose. (2006). *Data Mining Methods And Models.* Canada: John Wiley & Sons, Inc.

Moertini, V. (2002). *Data Mining Sebagai Solusi Bisnis.* Integral Vol. 7 No. 1.

PU, D. (2011). *Data Jalan & Jembatan Selama Pelita IV.* Departemen Pekerjaan Umum Dirjen Bina Marga.

Simatupang, J. E. (2011). *Studi Penentuan Prioritas Peningkatan Ruas Jalan Nasional Di Provinsi Kalimantan Selatan.* Malang: Brawijaya.

Suwandi, A. (2011). *Evaluasi Tingkat Kerusakan Jalan Dengan Metode Pavement Condition Index (PCI) Untuk Menunjang Pengambilan Keputusan.* Yogyakarta: UGM.

Syafarina, G. A. (2016). Penerapan Algoritma Neural Network Dalam Menentukan Prioritas Pengembangan Jalan di Provinsi Kalimantan Selatan. Banjarmasin: Jurnal Ilmiah Technologia