Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

Team project ©2017
Dony Pratidana S. Hum | Bima Agus Setyawan S. IIP

BAB III

METODOLOGI PENELITIAN

3.1 Gambaran Umum Objek Penelitian

Objek yang diteliti yaitu kemacetan yang terjadi pada ruas jalan utama dan ruas jalan cabang. Ruas jalan utama yaitu suatu ruas jalan besar yang akan menjadi acuan dalam memprediksi level kemacetan ruas jalan cabang. Sedangkan ruas jalan cabang yaitu ruas jalan yang terhubung langsung dengan ruas jalan utama.

Tabel 3.1 Ruas yang Menjadi Objek Penelitian

No	Nama Ruas Utama	Nama Ruas Cabang
1	Tol Jakarta-Tangerang (Arah Barat)	Dari Tol S. Parman
		Dari Jl. Tomang Raya
		Ke Tol Lingkar Luar (Arah Utara - Meruya 2)
		Ke Tol Lingkar Luar (Arah Selatan - Meruya 3)
2	Tol Jakarta-Tangerang (Arah Timur)	Dari Tol Jakarta-Tangerang (Arah Timur - Sebelum)
		Dari Tol Lingkar Luar (Arah Selatan)
		Ke Jl. Letjen. S. Parman
		Ke Jl. Tomang Raya
3	Tol Lingkar Luar Jakarta (Arah Utara)	Dari Tol Jakarta-Tangerang (Arah Barat)
		Dari Tol Lingkar Luar Jakarta (Arah Utara - Sebelum)
		Ke Tol Prof Sedyatmo (Arah Timur)
		Ke Tol Prof Sedyatmo (Arah Barat)
4	Tol Lingkar Luar Jakarta (Arah Selatan)	Dari Tol Lingkar Luar Jakarta (Arah Selatan - Sebelum)
		Dari Jl. Tol Airport Prof Sedyatmo
		Ke Tol Lingkar Luar Jakarta (Arah Selatan - Selanjutnya)
		Ke Tol Jakarta-Tangerang (Arah Timur)

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

Masing-masing ruas utama dilakukan pengumpulan data terhadap level kemacetan yang terjadi pada radius 2 Km setelah masuk tol dan 2 Km sebelum keluar tol. Sedangkan untuk ruas cabang dilakukan pengumpulan data terhadap level kemacetan yang terjadi pada radius 500 meter. Ruas utama dan ruas cabang yang dijadikan sebagai objek penelitian dapat dilihat pada tabel 3.1.

3.2 Penelitian Sebelumnya

Sebelum penelitian ini dilakukan, diperlukan studi terhadap penelitian-penelitian sebelumnya. Penelitian pertama yaitu "Perancangan Sistem Pencarian Rute Alternatif Di Bandung Untuk Menghindari Kemacetan Lalu Lintas Dengan Memanfaatkan Google *Aplication Programming Interface* (API) Berbasis Android" (Sari, 2014). Penelitian ini bertujuan untuk membantu pengguna jalan untuk memberikan informasi mengenai kondisi kemacetan lalu-lintas pada beberapa rute alternatif. Penelitian menggunakan bantuan Google Traffic untuk mendapatkan informasi mengenai kondisi kemacetan jalan dengan algoritma Djikstra untuk mendapatkan rute-rute alternatif.

Penelitian kedua berjudul “Analisa dan Penerapan Metode C4.5 Untuk Prediksi Loyalitas Pelanggan” (Santoso, 2012). Penelitian tersebut bertujuan untuk memprediksi loyalitas pelanggan pada suatu perusahaan telekomunikasi. Dalam mencari pola loyalitas pelanggan, digunakan algoritma C4.5. Dalam melakukan penelitian tersebut, digunakan sebuah *software Data Mining* yang bernama WEKA *Classifier*. Data sampel yang

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

digunakan ialah data primer yang diambil pada Oktober sampai dengan November 2013 sebanyak 40 data. Tingkat akurasi dari prediksi yang dihasilkan mencapai 97.5%.

Penelitian ketiga berjudul "*Prediction of Warning Level in Aircraft Accidents Using Data Mining Techniques*" (Christopher, 2016). Penelitian tersebut bertujuan untuk menemukan pola yang mempengaruhi *warning level* kecelakaan pada pesawat terbang. Ada beberapa metode yang digunakan, yaitu *Decision Tree* (C4.5), *Naive Bayes*, *Support Vector Machine*, *K-Nearest Neighbour*, dan *Neural Network*. Data sampel yang digunakan ialah data dari *database* sebuah perusahaan penerbangan dari tahun 1970 sampai dengan 2012.

Classifiers	Accuracy (%)
Decision tree	97·68
Naive Bayes	97·28
Support vector machines	97·65
K-nearest neighbour	96·83
Neural network	95·05

Gambar 3.1 Perbandingan Akurasi Algritma Klasifikasi

Gambar 3.1 merupakan perbandingan akurasi dari algoritma-algoritma yang digunakan pada penelitian tersebut. Yang mana algoritma *decision tree* (C4.5) memiliki akurasi tertinggi yaitu 97.68%, dan disusul oleh *Naive Bayes* yaitu 97.28%.

Sedangkan ringkasan penelitian terdahulu dapat dilihat pada tabel 3.2 di bawah ini.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

Tabel 3.2 Penelitian Terdahulu

No	Peneliti	Permasalahan	Solusi	Kesimpulan
1.	Pratiwi Widhi Maya Sari, Gelar Budiman, Ratri Dwi Atmaja	Merancang sebuah sistem yang dapat memberikan rute alternatif untuk menghindari kemacetan lalu lintas di Bandung.	Membaca nilai RGB pada Google Traffic untuk mengetahui level kemacetan dan menggunakan algoritma Djikstra.	Dengan membaca warna RGB dari Google Traffic level kemacetan suatu ruas jalan dapat diketahui. Rute-rute alternatif yang dihasikan oleh algoritma Djikstra disajikan dengan waktu tempuh masing-masing.
2.	Teguh Budi Santoso	Memprediksi loyalitas pelanggan pada sebuah perusahaan telekomunikasi.	Menggunakan algortima C4.5 diterapkan pada perangkat lunak WEKA untuk memprediksi loyalitas pelanggan.	Algoritma C4.5 dapat digunakan untuk data *training* yang cukup terbatas, dan memberikan akurasi tinggi, yaitu 97.5%.
3.	A. B. Arockia Christopher, S. Appavu	Menemukan pola yang mempengaruhi *warning level* kecelakaan pada pesawat terbang, dan membandingkan tingkat akurasi dari algoritma-algoritma yang dipakai.	Menggunakan algoritma *Decision Tree* (C4.5), *Naive Bayes*, *Support Vector Machine*, *K-Nearest Neighbour*, dan *Neural Network* untuk menemukan pola yang mempengaruhi *warning level* kecelakaan pada pesawat terbang.	Algoritma C4.5 memberikan akurasi yang terbaik dibandingkan algoritma klasifikasi lainnya.

3.3 Metode Penelitian

3.3.1 Algoritma C4.5

Dalam penelitian ini, penulis memutuskan untuk menggunakan metode algoritma C4.5. Algoritma C4.5 penulis pertimbangkan untuk digunakan pada penelitian ini karena mudah dipahami, dan pada beberapa penelitian sebelumnya algoritma tersebut telah memberikan akurasi yang cukup tinggi dibandingkan algoritma klasifikasi lainnya. Tabel 3.3 merupakan perbandingan akurasi antara algoritma C4.5 dan *Naive Bayes* yang telah penulis ringkas dari beberapa penelitian sebelumnya.

Tabel 3.3 Perbandingan Akurasi Algoritma C4.5 dan Naive Bayes

Judul Penelitian	Penulis	Tahun	Akurasi (%)		Teknik Pengujian
			C4.5	Naive Bayes	
Analisis Perbandingan Algoritma *Classification* untuk *Authentication* Uang Kertas (Studi Kasus: *Banknote Authentication*)	Khairul Sani, Wing Wahyu Winarno, Silmi Fauziati	2016	98,5	85	*Cross Validation*
Perbandingan Kinerja Metode Klasifikasi *Data Mining* Menggunakan Naive Bayes dan Algoritma C4.5 untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa	Fiastantyo Gian	2014	77,354	74,094	Sampling Linear
Analisa Perbandingan Algoritma SVM,	Dwi Widiastuti	2008	97,761	93,519	*Cross Validation*

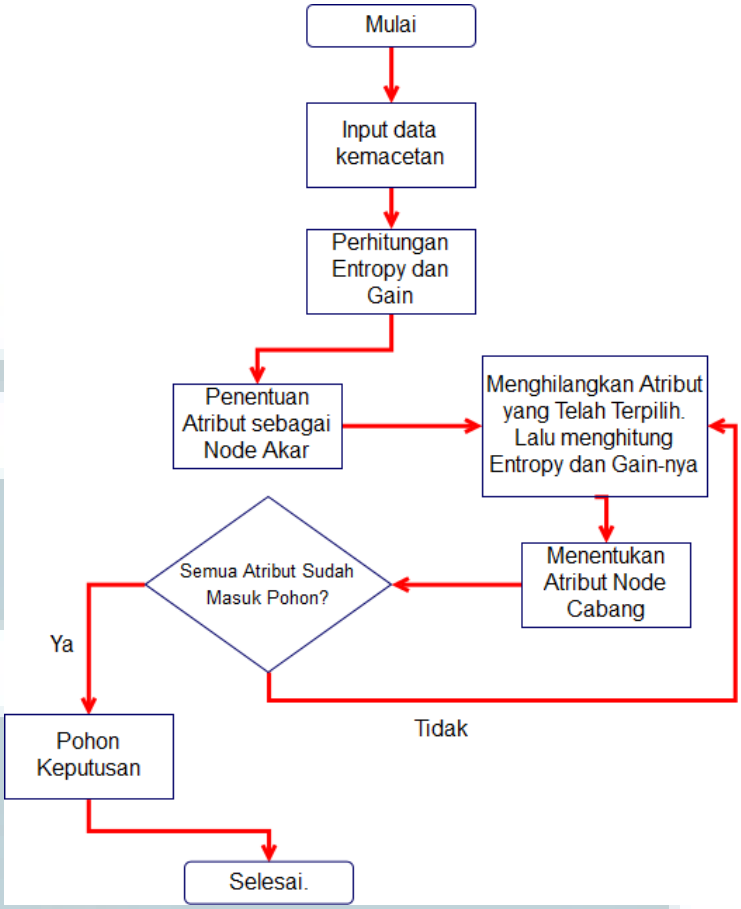
Judul Penelitian	Penulis	Tahun	Akurasi (%)		Teknik Pengujian
			C4.5	Naive Bayes	
Naive Bayes, dan Decision Tree Dalam Mengklasifikasikan Serangan (Attacks) pada Sistem Pendeteksi Intrusi					

Dengan demikian, pada penelitian ini algoritma C4.5 digunakan untuk mencari pola kemacetan yang terjadi berdasarkan data-data yang telah dikumpulkan selama beberapa waktu. Hasil akhir dari algoritma ini adalah sebuah pohon keputusan yang berisikan *rule* untuk memprediksi kemacetan.

Terdapat dua buah *dataset* yang akan digunakan pada algoritma ini. *Dataset* tersebut yaitu *dataset* ruas utama, dan *dataset* ruas cabang. Sehingga nantinya algoritma tersebut digunakan untuk menghasilkan dua buah *rule*, yaitu *rule* kemacetan ruas utama dan *rule* kemacetan ruas cabang.

Dalam penerapannya, dalam membentuk pohon keputusan algoritma C4.5 memiliki tahapan-tahapan yang harus dilakukan. Tahapan-tahapan tersebut seperti *flowchart* pada gambar 3.2.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017



Gambar 3.2 *Flowchart* Algoritma C4.5

Keterangan:

1. Proses melakukan *input* data kemacetan dari *dataset* ruas utama dan ruas cabang.
2. Proses perhitungan *Entropy* dari tiap-tiap atribut dan dilanjutkan perhitungan *Gain*.
3. Proses penentuan atribut sebagai akar. Atribut dengan *Gain* tertingi akan dijadikan sebagai akar.
4. Proses menghilangkan atribut yang telah terpilih sebagai akar. Kemudian dilakukan perhitungan *Entropy* dan *Gain*.

5. Proses menentukan *Node* cabang, yaitu memilih atribut dengan nilai *Gain* tertinggi.
6. Melakukan pemeriksaan apakah semua atribut telah masuk ke dalam pohon. Jika sudah, lanjutkan ke tahap selanjutnya. Jika belum, kembali pada tahap perhitungan *Entropy* dan *Gain* dengan menghilangkan atribut yang telah terpilih.
7. Hasil dari pohon keputusan.

Algoritma C4.5 akan digunakan untuk perhitungan dua buah *dataset*, yaitu *dataset* ruas utama dan ruas cabang. Masing-masing *dataset* memiliki atributnya masing-masing untuk digunakan pada algoritma C4.5. Atribut yang akan digunakan dapat dilihat pada tabel 3.4.

Tabel 3.4 Atribut yang Digunakan

***Dataset* Ruas Utama**	***Dataset* Ruas Cabang**
Atribut	
1. Nama Ruas Utama 2. Area 3. Hari 4. Jam 5. Kondisi (Kelas)	1. Nama Ruas Utama 2. Area Ruas Utama 3. Kondisi Ruas Utama 4. Nama Ruas Cabang 5. Hari 6. Jam 7. Kondisi (Kelas)

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

3.3.2 *Rapid Application Development*

Pada penelitian ini, model yang digunakan oleh penulis untuk pengembangan sistem yang akan dibuat adalah model *Rapid Application Development* (RAD). Model RAD dipakai oleh penulis karena lebih cocok digunakan dalam pengembangan aplikasi prediksi kemacetan ini. Tabel 3.5 merupakan perbandingan antara metode *Waterfall* dan RAD.

Tabel 3.5 Perbandingan Model *Waterfall* dan RAD

Model *Waterfall*	Model RAD
• Lebih cocok digunakan untuk pekerjaan yang memiliki ruang lingkup besar	• Lebih cocok digunakan untuk pekerjaan yang memiliki ruang lingkup kecil dan besar
• Tahapan yang dilalui lebih kompleks dan kaku, karena berbentuk *sequence*. Sehingga apabila ada kesalahan dalam suatu tahap, harus mengulang tahapan dari awal.	• Tahapan yang dilalui lebih sederhana dan cenderung lebih fleksibel. Tahapan dalam *Workshop design* dapat dilakukan berulang-ulang, hingga konsep dari sistem telah didefinisikan dengan baik.
• Cenderung memakan waktu lebih lama.	• Cenderung memakan waktu lebih cepat.
• Berfokus pada kualitas sistem.	• Berfokus pada waktu penyelesaian proyek.
• Biaya yang dikeluarkan cenderung lebih besar.	• Biaya yang dikeluarkan cenderung lebih kecil.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

Gambar 3.3 merupakan tahapan dalam pengembangan aplikasi prediksi kemacetan dengan menerapkan model *Rapid Application Development* (RAD) :

Gambar 3.3 Tahapan RAD

1. *Requirements Planning* (Perencanaan Syarat-Syarat)

   Pada tahap awal ini dilakukan pengumpulan data yang dibutuhkan oleh sistem, yaitu data-data kemacetan pada beberapa ruas jalan di wilayah Jakarta Barat.

   Sebelum pengumpulan data dilakukan, penulis melakukan penentuan titik-titik koordinat pengambilan data ruas jalan, serta dilakukan pembangunan aplikasi yang dapat mengambil data mentah kemacetan dari waktu ke waktu secara otomatis melalui Google Maps. Selanjutnya aplikasi penarikan data tersebut dijalankan selama 6 minggu. Setelah mencapai waktu yang ditentukan, data mentah akan dilakukan *preprocessing* untuk menghasilkan data final kemacetan ruas jalan.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

2. RAD *Design Workshop* (*Workshop* Desain RAD)

Pada tahap ini dilakukan perancangan aplikasi prediksi kemacetan. Pada perancangan sistem ini, penulis menggunakan *Unified Modeling Language* berupa *Use Case Diagram*, *Activity Diagram*, dan *Class Diagram*.

Selanjutnya dilakukan penerapan rancangan sistem yang telah dibuat ke dalam pembangunan aplikasi (*coding*). Pembangunan aplikasi sepenuhnya mengacu pada rancangan aplikasi yang telah dibuat sebelumnya.

3. *Implementation* (Implementasi)

Pada tahap ini sistem yang dimaksud telah dihasilkan. Selanjutnya dilakukan *testing* terhadap sistem yang telah dibuat. Dalam aplikasi ini *testing* yang dilakukan ialah untuk mengukur tingkat akurasi dari prediksi kemacetan yang dihasilkan.

Selanjutnya dilakukan *upload* data-data *server side* dari *localhost* ke dalam *server* yang telah disewa sebelumnya.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

3.4 Variabel Penelitian

3.4.1 Variabel Independen

Variabel independen merupakan variabel yang menjelaskan atau mempengaruhi variabel lain. Pada penelitian ini, variabel independen yang digunakan adalah level kemacetan ruas jalan utama jalur masuk (X_1) dan jalur keluar (X_2). Serta level kemacetan ruas jalan cabang (X_3).

3.4.2 Variabel Dependen

Variabel dependen merupakan variabel yang dijelaskan atau dipengaruhi oleh variabel independen. Pada penelitian ini, variabel dependen yang digunakan untuk prediksi level kemacetan ruas utama adalah Nama Ruas Jalan Utama (Y_1), Hari (Y_2), Jam (Y_3).

Sedangkan untuk prediksi level kemacetan pada ruas jalan cabang variabel yang digunakan ialah nama Ruas Jalan Cabang (Y_4), Jenis Akses ke Ruas Utama (Y_5), Level kemacetan ruas utama (Y_6).

3.5 Teknik Pengumpulan Data

Pengumpulan data dilakukan dengan cara observasi. Observasi dilakukan dengan mengamati kemacetan yang terjadi pada ruas jalan Tol Jakarta-Tangerang dan Tol Lingkar Luar Jakarta berserta masing-masing ruas cabangnya melalui Google Traffic melalui aplikasi penarikan data otomatis yang akan dibuat. Aplikasi tersebut berjalan secara otomatis dengan jeda waktu 1 jam selama 6 minggu, terhitung mulai dari 12 April 2017 pukul 00:00 WIB sampai dengan 23 Mei 2017 pukul 23:00 WIB.

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017

Waktu selama 6 minggu dipilih karena keterbatasan sumber daya yang digunakan.

3.6 Teknik Simulasi Data

Pada penelitian ini, penyimpanan data mentah dan final kemacetan ruas jalan menggunakan DBMS MySQL. Algoritma C4.5 diterapkan pada *software Data Mining* WEKA untuk menemukan pola kemacetan berdasarkan data-data kemacetan yang ada di dalam *database*. Kemudian *rule* yang telah dihasilkan oleh WEKA, diterapkan pada aplikasi Android yang dibuat untuk memprediksi kemacetan yang terjadi pada ruas jalan utama dan dampak kemacetan ruas jalan utama terhadap ruas jalan cabang. Tabel 3.4 merupakan perbandingan *tool* RapidMiner dan WEKA.

Gambar 3.4 Perbandingan *Tools* RapidMiner dan WEKA

RapidMiner	WEKA
Gratis untuk data dibawah 10.000 baris	Gratis
Perlunya melakukan perancangan *operator* sebelum melakukan pengolahan data	Penggunaan sederhana untuk melakukan pengolahan data
Memakan banyak memori, sehingga cukup berat saat digunakan	Sedikit memakan memori, sehingga cukup ringan saat digunakan
User Interface menarik	*User Interface* kurang menarik
Tools sangat lengkap	*Tools* cukup lengkap

PEmodelan Prediksi Kemacetan..., Dimas Dwi Cahyo, FTI UMN, 2017