

KAJIAN AWAL PROFILING KEJAHATAN DAN STRATEGI DALAM USAHA MENCEGAH TERJADINYA TINDAK KRIMINALITAS DI KABUPATEN SLEMAN

Agus Hindarto Wibowo^{*1)}, Andi Rahadiyan Wijaya²⁾

^{1, 2, 3)}Program Studi Teknik Industri, Fakultas Teknik
Universitas Gadjah Mada

Jalan Grafika No.2, Kampus UGM, Yogyakarta 55281

Telp. 085641246300

Email: bagushindartowibowo@yahoo.com

ABSTRAK

Kejahatan adalah tindakan yang melanggar hukum dan dapat merugikan orang lain. Seiring berjalannya waktu, kejahatan meningkat dengan pola yang tidak menentu, Kejahatan secara klasik "tidak dapat diprediksi". Ini tidak harus acak, tapi juga tidak dilakukan secara konsisten dalam ruang atau waktu, oleh karena itu perlu untuk mencegah tindak pidana yang akan terjadi. Data mining adalah teknik yang bisa digunakan untuk mencegah kejahatan. Menurut penelitian sebelumnya, teknik data mining memiliki berbagai metode yang dapat digunakan untuk memprediksi kejahatan dengan memanfaatkan data yang ada. Pada penelitian ini, penulis berusaha untuk menambang data (*data mining*) laporan tindak kriminal yang terjadi di Kabupaten Sleman dari tahun sebelumnya. Data dianalisis menggunakan algoritma *decision tree* yang mana algoritma ini bertujuan untuk melakukan klasifikasi data pada kelas tertentu, yang kemudian pola tersebut dapat digunakan untuk memperkirakan kejahatan yang akan terjadi. Hasil yang didapat adalah ketidaksesuaian antara data laporan kriminal pencurian dan penipuan terhadap rule yang diperoleh menggunakan algoritma *decision tree* pada *software Rapidminer* yang ditunjukkan oleh akurasi sebesar 40,00%. Nilai ini menunjukkan bahwa sebanyak 40,00% dari 50 data *testing set* terdapat kesesuaian prediksi antara data laporan kriminal dengan *rule* yang diperoleh menggunakan algoritma *decision tree*.

Kata kunci: *Classification, Clustering, Crime profiling, Data mining, Predict, Decision tree.*

1. Pendahuluan

Kejahatan seiring berjalannya waktu semakin berkembang dengan cepat, terutama di Indonesia. Kejahatan merupakan tindakan yang melanggar aturan hukum yang merugikan orang lain dan melanggar norma-norma yang ada di masyarakat. Kejahatan adalah masalah sosial-ekonomi yang mempengaruhi orang-orang di seluruh dunia dan berdampak negatif pada kesejahteraan masyarakat. (Pereira dan brandao., 2014). Kemudian kejahatan adalah gangguan sosial dan dapat merugikan banyak masyarakat dengan berbagai macam cara (Kaur dan Singh., 2017). Kejahatan tidak sistematis atau sama sekali acak atau tidak dapat diprediksi secara langsung (Yu dkk., 2011). Penelitian tentang penggunaan *data mining* untuk memprediksi kejahatan juga pernah dilakukan oleh Shojaee dkk (2014), Tayal dkk (2014), Zubi dan Mahmud (2013) yang memprediksi tingkat kriminalitas dengan tujuan untuk membantu pihak kepolisian dalam mencegah tindak kriminal yang akan terjadi, penelitian ini dilakukan karena tingkat kriminalitas yang terus meningkat dari tahun ke tahun.

Oleh karena itu, perlu dilakukan sebuah tindakan pencegahan yang dapat meminimalisir tingkat kejahatan yang akan terjadi. Salah satu teknik yang dapat digunakan untuk mencegah kejahatan yang akan terjadi adalah teknik *data mining* dengan cara memprediksi pola dari tindak kriminal yang akan terjadi berdasarkan data sebelumnya yang sudah ada. *Data mining* adalah teknik yang memungkinkan untuk memprediksi kejahatan yang akan terjadi secara akurat (Yu dkk., 2011). Melakukan penelitian dengan teknik *data mining* untuk mendeteksi dan memprediksi pola kejahatan dapat membantu memecahkan permasalahan kejahatan lebih cepat (Nath., 2006).

2. Metode

2.1 Data mining

Data mining adalah suatu teknik pengumpulan data yang didapat dari berbagai sumber, yang kemudian dari data-data tersebut diubah menjadi informasi yang sangat berguna menggunakan berbagai metode yang sudah ditentukan. *Data mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk penanganan permasalahan pengambilan informasi dari *database* yang besar (Larose., 2014). Secara garis besar, *data mining* dapat dikelompokkan menjadi 2 kategori utama, yaitu *descriptive mining* dan *predictive* (Tan dkk., 2005). Dalam melakukan *predictive*, diperlukan teknik pengolahan data yang sudah ada sebelumnya, yang kemudian dikumpulkan dan diolah, teknik yang paling populer digunakan adalah *Data Mining* (Liao dkk., 2012). Beberapa metode di dalam teknik *data mining* telah digunakan untuk memprediksi sesuatu pola, dalam hal ini adalah memprediksi pola tindak kejahatan. Menurut Liao dkk (2012), *data mining* adalah teknik yang paling populer digunakan dalam kurun waktu sepuluh tahun terakhir dari tahun 2000 hingga tahun 2011. Informasi didalam *data mining* memiliki tipe yang berbeda (Kaur and Singh., 2017).

2.2 Klasifikasi

Klasifikasi *data mining* adalah penempatan objek-objek ke salah satu dari beberapa kategori yang telah ditentukan sebelumnya. Klasifikasi banyak digunakan untuk memprediksi kelas pada suatu label tertentu, yaitu dengan mengklasifikasikan data (membangun model) berdasarkan *training set* dan nilai-nilai (label kelas) dalam mengklasifikasikan atribut tertentu dan menggunakannya dalam mengklasifikasikan data yang baru (*testing set*) (Susanto & Sudiyatno, 2014).

Data yang digunakan sebagai input adalah 100 data laporan kriminal mengenai delapan variabel yang dianggap memiliki pengaruh terhadap terjadinya tindak kejahatan pencurian dan penipuan. Dari 100 data laporan kriminal tersebut akan dibagi menjadi dua macam data, yaitu *training set* dan *testing set*. Banyak penelitian terdahulu yang telah menjelaskan rasio yang digunakan dalam menentukan *training set* dan *testing set*. Sebanyak 60% dari keseluruhan data digunakan untuk data *training* yang nantinya akan menghasilkan sebuah *rule*. Sisanya sebanyak 40% digunakan sebagai data *testing* (Sulistiyowati, 2016).

2.3 Algoritma decision tree

Menurut Jailani dkk 2015, *Decision Tree* (pohon keputusan) adalah pohon yang ada dalam analisis pemecahan masalah, pemetaan mengenai alternatif-alternatif pemecah masalah yang dapat diambil dari masalah. Pohon keputusan juga dapat disebut salah satu algoritma dalam klasifikasi yang paling populer karena mudah untuk diinterpretasi. Konsep dari *decision tree* adalah mengubah data menjadi pohon keputusan dan aturan-aturan keputusan (*rule*). *Decision tree* sesuai digunakan untuk kasus dengan *output* berupa nilai diskrit. Manfaat utama dari penggunaan *decision tree* adalah kemampuannya untuk menelaah dan menguraikan proses pengambilan keputusan yang kompleks sehingga menjadi lebih simpel dan pengambil keputusan lebih mudah dalam interpretasikan solusi dari sebuah masalah.

Pohon keputusan biasanya digunakan untuk mendapatkan informasi untuk tujuan pengambilan sebuah keputusan. Pohon keputusan dimulai dengan sebuah *root node* (titik awal) yang digunakan oleh user untuk mengambil tindakan. Berdasarkan *node root* ini, user memecahkan *leaf node* sesuai dengan algoritma *decision tree*. Hasil akhir dari penyusunan *node root* dan *leaf node* tersebut adalah sebuah pohon keputusan dengan setiap cabangnya menunjukkan kemungkinan skenario dari keputusan yang diambil serta hasilnya. Konsep pohon

keputusan adalah mengubah data menjadi sebuah pohon keputusan (*decision tree*) dan aturan-aturan keputusan (*rule*) (Susanto & Sudiyatno., 2014).

2.4 Objek penelitian

Penelitian ini dilakukan dengan tujuan untuk meninjau dan mengklasifikasikan atribut atau variabel yang dapat mempengaruhi terjadinya tindak kriminalitas pencurian dan penipuan. Metode yang digunakan dalam penelitian ini adalah teknik klasifikasi *data mining* dengan metode algoritma *decision tree*. Hasil penelitian yang ingin dicapai adalah *rule* atau aturan dari kejahatan pencurian dan penipuan berdasarkan atribut atau variabel yang ada. Pengolahan data dilakukan dengan bantuan *software* Ms. Excel, SPSS, dan *Rapidminer*. Pada penelitian pendahuluan ini terdapat 100 data laporan kriminal pencurian dan penipuan di Kabupaten Sleman pada tahun 2017 yang dijadikan objek dalam penelitian ini.

2.5 Metode pengumpulan data

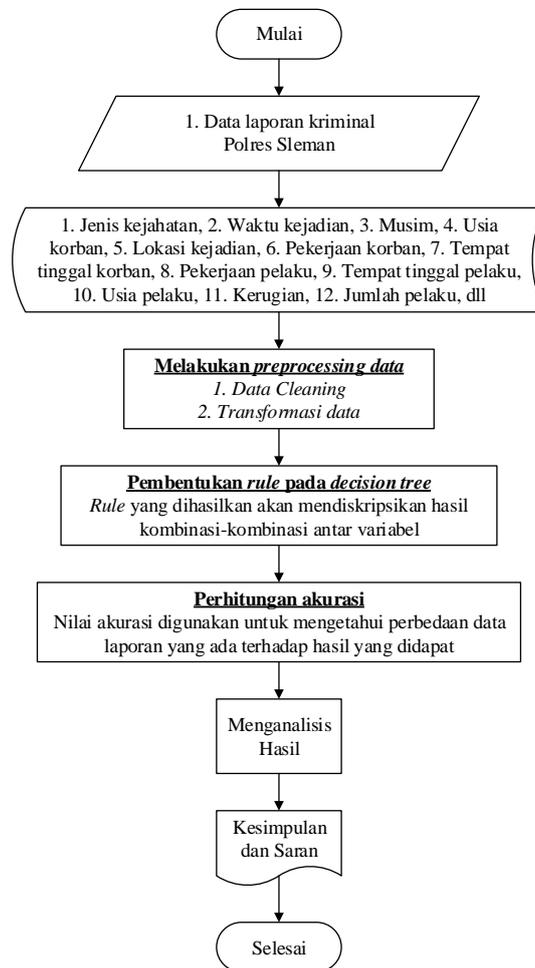
Menurut (Taufik dkk., 2012), penelitian kuantitatif adalah penelitian yang dimaksudkan untuk memperoleh data yang berbentuk angka atau data kualitatif yang diangkakan. Pada penelitian pendahuluan ini, data yang diperoleh merupakan data atribut atau variabel yang mempengaruhi terjadinya kriminalitas pencurian dan penipuan yaitu hari, musim, waktu, TKP, jenis kelamin korban, usia korban, jenis kelamin pelaku, dan usia pelaku.

2.6 Jenis data

Data sekunder adalah data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian dari sumber-sumber yang telah ada (Viarani & Zadry., 2015). Pada penelitian pendahuluan ini, data sekunder yang digunakan merupakan 100 data atribut atau variabel dari laporan tindak kriminal pencurian dan penipuan yaitu hari, musim, waktu, TKP, jenis kelamin korban, usia korban, jenis kelamin pelaku, dan usia pelaku. Data sekunder lain dalam penelitian ini adalah data penelitian terdahulu mengenai masalah terkait yaitu teknik klasifikasi data mining, metode algoritma *decision tree* dengan permasalahan terkait kriminalitas.

2.7 Alur penelitian

Gambar 1 menunjukkan gambar diagram alir dari keseluruhan tahap penelitian yang dilakukan.



Gambar 1. Diagram alur penelitian

3. Hasil dan Pembahasan

Tabel 1 menunjukkan 5 data dari 100 data laporan kriminal yang telah dilakukan *data cleaning*.

Tabel 1. Hasil data *cleaning*

D a t a	Hari	Musim	Waktu	TKP	Jenis kelamin korban	Usia korban	Jenis kelamin Pelaku	usia pelaku	Jenis kejahatan
1	weekday	Penghujan	pagi	Pnginapan	P	Lansia	P	dewasa	Pencurian
2	weekday	Penghujan	siang	Pnginapan	P	Dewasa	L	dewasa	Pencurian
3	weekday	Penghujan	pagi	kantor	L	Manula	L	dewasa	Pencurian
4	weekend	Penghujan	Siang	toko	L	Dewasa	L	remaja	Pencurian
5	weekend	Penghujan	siang	rumah	L	Dewasa	L	dewasa	Pencurian

Data yang telah dibersihkan dan diseleksi selanjutnya dilakukan transformasi data. Proses data transformasi adalah proses untuk mengubah data menjadi format yang sesuai dalam proses

pengolahan *data mining*. Setelah itu sebelum dilakukannya pengolahan data untuk memperoleh *rule* pada *decision tree*, peneliti perlu membagi data menjadi dua bagian, yaitu *training set* dan *testing set*. Rasio pembagian antara *training set* dan *testing set* dilakukan berdasarkan aturan-aturan yang umum atau berdasarkan penelitian sebelumnya. Sulistyowati, 2016 menjelaskan sebanyak 60% dari keseluruhan data digunakan untuk data *training* yang nantinya akan menghasilkan sebuah *rule*. Sisanya sebanyak 40% digunakan sebagai data *testing*. Tabel 2 menunjukkan aturan transformasi dan tabel 3 menunjukkan hasil data transformasi.

Tabel 2. Aturan transformasi

Atribut/variabel	Kode	Penjelasan
Hari	1 dan 2	Weekday dan weekend
Musim	1 dan 2	Penghujan dan kemarau
Waktu	1,2,3,4,5	Pagi,siang,sore,malam,dini hari
TKP	1,2,3,4,5	Rumah,kantor,toko,penginapan,umum
Jenis kelamin korban	1 dan 2	L dan P
Usia korban	1,2,3,4	Remaja,dewasa,lansia,manula
Jenis kelamin pelaku	1 dan 2	L dan P
Usia pelaku	1,2,3,4	Remaja,dewasa,lansia,manula

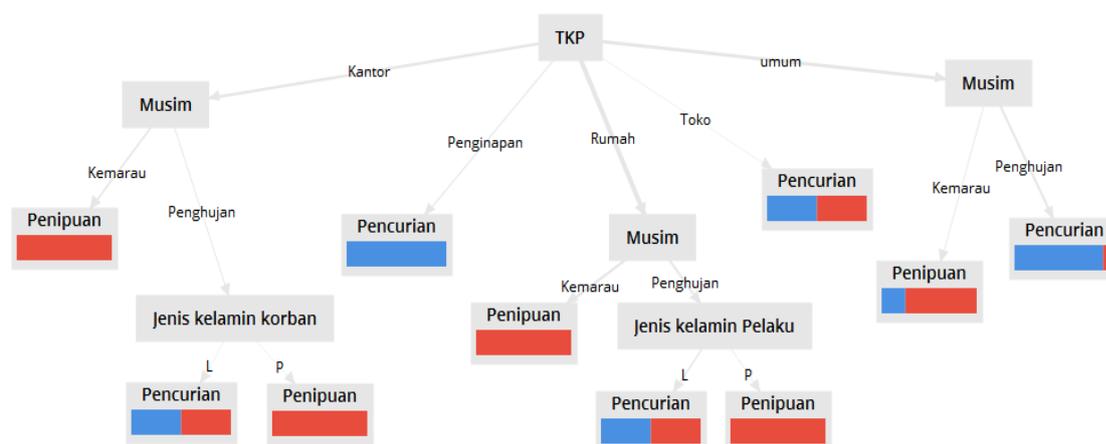
Tabel 3. Hasil data transformasi

Data	Hari	Musim	Waktu	TKP	Jenis kelamin korban	Usia korban	Jenis kelamin Pelaku	usia pelaku	Jenis kejahatan
1	1	1	1	4	2	3	2	2	Pencurian
2	1	1	2	4	2	2	1	2	Pencurian
3	1	1	1	2	1	4	1	2	Pencurian
4	2	1	2	3	1	2	1	1	Pencurian
5	2	1	2	1	1	2	1	2	Pencurian

Bentuk uji hipotesis dari setiap variabel adalah H_0 = Semua proporsi setiap variabel terhadap jenis kejahatan adalah tidak terdapat pengaruh variabel terhadap jenis kejahatan. H_a = Tidak semua proporsi setiap variabel terhadap jenis kejahatan adalah tidak sama atau terdapat pengaruh variabel terhadap jenis kejahatan pencurian dan penipuan.

Hasil dan analisis Uji Independensi (*Chisquare*) menggunakan *software* SPSS, jika setiap variabel menunjukkan nilai $sig = < 0,05$, maka H_0 ditolak. Pada kasus penelitian ini variabel hari, TKP, usia korban, dan usia pelaku menunjukkan nilai $sig > 0,05$, sehingga variabel-variabel tersebut dianggap tidak terdapat pengaruh terhadap jenis kejahatan pencurian dan penipuan. Kemudian variabel tersisa yang terdapat pengaruh, digunakan untuk melakukan analisis klasifikasi menggunakan algoritma *decision tree* pada *software* rapidminer.

Berdasarkan analisis klasifikasi menggunakan algoritma *decision tree* pada *Rapidminer* diperoleh *decision tree* sebagai berikut:



Gambar 2. Hasil *decision tree*

Rule yang terbentuk dari analisis klasifikasi pada permasalahan awal untuk *profiling* kejahatan pencurian atau penipuan di Kabupaten Sleman adalah sebagai berikut:

- IF** Tempat Kejadian Perkara = di Kantor dan
 - | Musim = Kemarau: **THEN** terjadi Penipuan {Pencurian=0, Penipuan=7}
 - | Musim = Penghujan dan
 - | | Jenis kelamin korban = Laki-laki: **THEN** terjadi Pencurian {Pencurian=1, Penipuan=1}
 - | | Jenis kelamin korban = Perempuan: **THEN** terjadi Penipuan {Pencurian=0, Penipuan=2}
- IF** Tempat Kejadian Perkara = di Penginapan: **THEN** terjadi Pencurian {Pencurian=5, Penipuan=0}
- IF** Tempat Kejadian Perkara = di Rumah dan
 - | Musim = Kemarau: **THEN** terjadi Penipuan {Pencurian=0, Penipuan=10}
 - | Musim = Penghujan dan
 - | | Jenis kelamin Pelaku = Laki-laki: **THEN** terjadi Pencurian {Pencurian=3, Penipuan=3}
 - | | Jenis kelamin Pelaku = Perempuan: **THEN** terjadi Penipuan {Pencurian=0, Penipuan=3}
- IF** Tempat Kejadian Perkara = di Toko: **THEN** terjadi Pencurian {Pencurian=1, Penipuan=1}
- IF** Tempat Kejadian Perkara = di tempat umum dan
 - | Musim = Kemarau: **THEN** terjadi Penipuan {Pencurian=1, Penipuan=3}
 - | Musim = Penghujan: **THEN** terjadi Pencurian {Pencurian=8, Penipuan=1}

Tabel 4. Performansi data

	True penipuan	True pencurian	Class precision
Prediksi penipuan	4	29	12,12%
Prediksi pencurian	1	16	94,12%
Class recall	80.00%	35.56%	

Accuracy: 40.00%

Dari tabel diatas, dapat terlihat bahwa ketidaksesuaian antara data laporan kriminal pencurian dan penipuan terhadap *rule* yang diperoleh menggunakan algoritma *decision tree* pada *software Rapidminer* yang ditunjukkan oleh akurasi sebesar 40,00%. Nilai ini menunjukkan bahwa sebanyak 40,00% dari 50 data *testing set* terdapat kesesuaian prediksi antara data laporan kriminal dengan *rule* yang diperoleh menggunakan algoritma *decision tree*.

Class prediction untuk prediksi penipuan sebesar 12,12% menunjukkan prediksi klasifikasi yang diperoleh dari perhitungan klasifikasi algoritma *decision tree* ini, kemudian ternyata sistem dapat melakukan pemanggilan ulang klasifikasi (*class recall*) untuk prediksi penipuan sebesar 80,00%. *Class prediction* untuk prediksi pencurian adalah sebesar 94,12% menunjukkan

prediksi klasifikasi yang diperoleh dari perhitungan klasifikasi algoritma *decision tree* ini, kemudian ternyata sistem dapat melakukan pemanggilan ulang klasifikasi (*class recall*) untuk pencurian sebesar 35,56%.

4. Simpulan

Berdasarkan hasil penelitian yang telah dilakukan maka didapat kesimpulan bahwa metode *decision tree* dapat menemukan pola kejahatan yang dapat dijadikan informasi berguna bagi pihak kepolisian dan masyarakat luas. *Rule* yang dihasilkan dapat menjadi informasi baru yang berguna dalam *profiling* kejahatan di Kabupaten Sleman. Penggunaan *software rapidminer* dapat menyajikan data berupa *tree* dan tingkat keakurasian kesesuaian antara data laporan dengan *rule* yang didapat menggunakan algoritma *decision tree*.

Setelah didapatkan hasil *rule* seperti di atas, diharapkan bisa dijadikan dasar untuk pihak kepolisian dalam usaha mencegah terjadinya tindak kriminal yang akan terjadi, selain itu hasil awal ini juga dapat untuk membantu masyarakat dalam meningkatkan tingkat kewaspadaan

Daftar Pustaka

- Jailani, Defit, S. & Nurcahyo, G. W., 2015. Penerapan Algoritma C4.5 pada NUPTK untuk Menentukan Pola Sertifikasi Guru dengan Menggunakan Metode Klasifikasi Decision Tree (Studi Kasus: Dinas Pendidikan Pemuda dan Olahraga Kabupaten Rokan Hulu). *Riau Journal of Computer Science*, I(1), pp. 69-83.
- Kaur, S., 2017. Systematic Review of Crime Data Mining. *International Journal of Advanced Research in Computer Science*, 8(5).
- Larose, D.T., 2014. *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons.
- Liao, S.H., Chu, P.H. and Hsiao, P.Y., 2012. Data mining techniques and applications—A decade review from 2000 to 2011. *Expert systems with applications*, 39(12), pp.11303-11311.
- Nath, S.V., 2006, December. Crime pattern detection using data mining. In *Web intelligence and intelligent agent technology workshops, 2006. wi-iat 2006 workshops. 2006 IEEE/WIC/ACM International Conference on* (pp. 41-44). IEEE.
- Pereira, B.L. and Brandão, W.C., 2014. ARCA: Mining crime patterns using association rules. *In IADIS International Conference Applied Computing*(Vol. 2014).
- Shojaee, S., Mustapha, A., Sidi, F. and Jabar, M.A., 2013. A study on classification learning algorithms to predict crime status. *International Journal of Digital Content Technology and its Applications*, 7(9), p.361.
- Sulistiyowati, 2016. Klasifikasi Data Kejadian Luar Biasa Campak menggunakan Metode Decision Tree C4.5. *Jurnal Informatika dan Komputer (JIKO)*, I(1), pp. 8-11.
- Susanto, H. & Sudiyatno, 2014. Data Mining untuk Memprediksi Prestasi Siswa berdasarkan Sosial Ekonomi, Motivasi, Kedisiplinan, dan Prestasi Masa Lalu. *Jurnal Pendidikan Vokasi*, IV(2), pp. 222-231.
- Tan, P.N., Steinbach, M. dan Kumar, V., 2006, *Introduction to Data Mining*, Pearson, San Fransisco.
- Taufik, R., Sumantri, Y. & Tantrika, C. F. M., 2012. *Penerapan Pemilihan Supplier Bahan Baku Ready Mix berdasarkan Integrasi Metode AHP dan Topsis (Studi Kasus pada PT. Merak Jaya Beton, Malang)*, Malang: Universitas Brawijaya.
- Tayal, D.K., Jain, A., Arora, S., Agarwal, S., Gupta, T. and Tyagi, N., 2015. Crime detection and criminal identification in India using data mining techniques. *AI & society*, 30(1), pp.117-127.
- Viarani, S. O. & Zadry, H. R., 2015. Analisis Pemilihan Pemasok dengan Metode Analytical Hierarchy Process di Proyek Indarung VI PT Semen Padang. *Jurnal Optimasi Sistem Industri*, pp. 55-70.
- Yu, C.H., Ward, M.W., Morabito, M. and Ding, W., 2011, December. Crime forecasting using data mining techniques. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on* (pp. 779-786). IEEE.

Zubi, Z.S. and Mahmud, A.A., 2014. Using Data Mining Techniques to Analyze Crime patterns in the Libyan National Crime Data. *Recent Advances in Image, Audio and Signal Processing*, 8, pp.79-85.