

Implementasi Metode Apriori Pada Sistem Persediaan Bahan Kimia Di Laboratorium Forensik Medan

Niken Aprilia, Mesran, Fince Tinus Waruwu

Program Studi Teknik Informatika STMIK Budi Darma Medan

Email: nikenoye@gmail.com

Abstrak—Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengestraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Dengan adanya data mining maka akan didapatkan suatu permata berupa pengetahuan di dalam kumpulan data-data yang banyak jumlahnya. Metode Apriori (RF) merupakan metode yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap node dilakukan secara acak. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari root node, internal node, dan leaf node dengan mengambil atribut dan data secara acak sesuai ketentuan yang diberlakukan. Root node merupakan simpul yang terletak paling atas, atau biasa disebut sebagai akar dari pohon keputusan. Solusi penyelesaikan dalam menentukan persediaan stock bahan kimia di Laboratorium Forensik Cabang Medan, dengan menerapkan metode Apriori untuk menentukan tingkat koefisien korelasi produk yang sering digunakan sehingga produk yang sering dibutuhkan dapat disediakan lebih banyak untuk menghindari kekosongan bahan kimia di Laboratorium Forensik Cabang Medan.

Kata Kunci: Data Mining; Metode Apriori; Bahan Kimia.

Abstract—Data mining is a term used to describe the discovery of knowledge in databases. Data mining is a process that uses statistical, mathematical, artificial intelligence, and machine learning techniques to extract and identify useful information and related knowledge from large databases. With data mining, a gem in the form of knowledge will be obtained in a large number of data sets. The Apriori (RF) method is a method that can improve accuracy results, because generating child nodes for each node is done randomly. This method is used to build a decision tree consisting of root nodes, internal nodes, and leaf nodes by taking attributes and data randomly according to the applicable provisions. The root node is the node located at the top, or commonly referred to as the root of the decision tree. The solution for determining chemical stock inventories at the Medan Branch Forensic Laboratory, by applying the Apriori method to determine the correlation coefficient level of frequently used products so that more frequently needed products can be provided to avoid chemical vacancies at the Medan Branch Forensic Laboratory.

Keywords: Data Mining; Apriori Method; Chemicals

1. PENDAHULUAN

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. *Datamining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengestraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Dengan adanya *data mining* maka akan didapatkan suatu permata berupa pengetahuan di dalam kumpulan data-data yang banyak jumlahnya [1].

Forensik berasal dari bahasa Latin yaitu, "forum" yang berarti tempat untuk melakukan transaksi. Pada perkembangan selanjutnya, forensik diperlukan pada pengungkapan suatu kasus tindak pidana dengan cara menyusun kembali (rekonstruksi) suatu tindak pidana itu dapat terjadi, sudah barang tentu berdasarkan bukti-bukti yang ada [2]. Ilmu Forensik dikategorikan ke dalam ilmu pengetahuan alam dan dibangun berdasarkan metode ilmu alam. Dalam pandangan ilmu alam, sesuatu dianggap ilmiah jika didasarkan pada fakta atau pengalaman (empirisme).

Permasalahan yang sering terjadi di Laboratorium Forensik Cabang Medan kurangnya bahan kimia yang dibutuhkan pada saat dilakukan penyelidikan, disebabkan sistem yang diterapkan kurang maksimal atau masih menggunakan sistem manual, sehingga bahan kimia yang *distock* di gudang terjadi penumpukan di beberapa item bahan kimia tetapi beberapa item yang lain kosong, sehingga menyebabkan proses penyelidikan menjadi terhambat karena bahan kimia yang dibutuhkan kosong dan harus menunggu dating bahan kimia yang dipesan, untuk menyelesaikan permasalahan tersebut maka dibuat suatu sistem data mining dengan metode *Apriori*, agar ke depannya bahan kimia di Laboratorium Forensik Cabang Medan tidak akan kekurangan *stock* [3].

Metode *Apriori* merupakan metode yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap *node* dilakukan secara acak. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari *root node*, *internal node*, dan *leaf node* dengan mengambil atribut dan data secara acak sesuai ketentuan yang diberlakukan. *Root node* merupakan simpul yang terletak paling atas, atau biasa disebut sebagai akar dari pohon keputusan. *Internal node* adalah simpul percabangan, dimana *node* ini mempunyai *output* minimal dua dan hanya ada satu *input*. Sedangkan *leaf node* atau *terminal node* merupakan simpul terakhir yang hanya memiliki satu *input* dan tidak mempunyai *output*. Pohon keputusan dimulai dengan cara menghitung nilai *entropy* sebagai penentu tingkat ketidakmurnian atribut dan nilai *information gain* [4].

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data *mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (*induction-based learning*) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari [5]. Data *mining* berisi pencarian trend atau pola yang diinginkan dalam database yang besar untuk membantu pengambilan keputusan di waktu yang akan datang. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lainnya[6], [7].

2.2 Apriori

Metode *Apriori* adalah pengembangan dari metode CART, yaitu dengan menerapkan metode *bootstrap aggregating (bagging)* dan *random feature selection* (Breiman 2001). Metode ini merupakan metode pohon gabungan. Dalam *Apriori*, banyak pohon ditumbuhkan sehingga terbentuk suatu hutan (*forest*), kemudian analisis dilakukan pada kumpulan pohon tersebut. Pada gugus data yang terdiriatasn amatan dan p peubah penjelas, prosedur untuk melakukan *Apriori* adalah [8]–[10]

1. Lakukan penarikan contoh acak berukuran n dengan pemulihan pada gugus data. Tahap ini adalah tahapan *bootstrap*.
2. Dengan menggunakan contoh *bootstrap*, pohon dibangun sampai mencapai ukuran maksimum (tanpa pemangkasan). Pembangunan pohon dilakukan dengan menerapkan *random feature selection* pada setiap proses pemilihan pemilah, yaitum peubah penjelas dipilih secara acak dimana $m < p$, lalu pemilah terbaik dipilih berdasarkan m peubah penjelas tersebut.
3. Ulangi langkah 1 dan 2 sebanyak k kali, sehingga terbentuk sebuah hutan yang terdiriatas k pohon.

Apriori memprediksi respons suatu amatan dengan cara menggabungkan (*aggregating*) hasil prediksik pohon. Untuk masalah klasifikasi, pohon yang dibangun adalah pohon klasifikasi dan hasil prediksi *Apriori* adalah berdasarkan *majority vote* (suaraterbanyak), yaitu kategori atau kelas yang paling sering muncul sebagai hasil prediksi darik pohon klasifikasi. Pada setiap iterasi *bootstrap*, terdapat sekitar sepertiga amatan gugus data asli yang tidak termuat dalam contoh *bootstrap*. Amatan tersebut disebut data *out-of-bag* (OOB). Data OOB tidak digunakan untuk membangun pohon, melainkan menjadi data validasi pada pohon yang bersesuaian. Nilai salah klasifikasi *Apriori* diduga melalui *error OOB* yang diperoleh dengan cara

1. Lakukan prediksi pada setiap poho yang besesuaian.
2. Umumnya setiap amatan gugus data asliakan menjadi data OOB sebanyak sekitar 36% atau sepertiga dari banyak pohon yang dibentuk. Oleh karena itu, pada langkah 1, masing-masing amatan gugus data asli mengalami prediksi sebanyak sekitar sepertiga kali dari banyaknya pohon. Jika a adalah sebuah amatan dari gugus data asli, maka hasil prediksi *Apriori* terhadap a adalah gabungan dari hasil prediksi setiap kali a menjadi data OOB.
3. *Error OOB* dihitung dari proporsimis klasifikasi hasil prediksi *Apriori* dari seluruh amatan gugus data asli.

Breiman dan Cutler (2005) menyarankanuntukmengamati error OOB saat k kecil, lalu memilih m yang menghasilkan error OOB terkecil. Jika *Apriori* dilakukan dengan menghasilkan variable importance, disarankan untuk menggunakan banyak pohon, misalnya 1000 pohon atau lebih. Jika peubah penjelas yang dianalisis sangat banyak, nilai tersebut dapat lebih besar agar *variable importance* yang dihasilkan semakin stabil. Dengan diterapkannya metode *Apriori* ini pada permasalahan data mining untuk prediksi *stock* bahan kimia yang paling dibutuhkan dan bahan kimia yang memangakan di *stock* berlebih sebagai cadangan persediaan, diharapkan dapat membantu dalam menyusun rencana kebutuhan bahan kimia dengan strategi yang tepat untuk menekan dan mengurangi penumpukan bahan kimia yang tidakterpakai.

Proses pembentukan C1 atau disebut dengan 1 itemset dengan jumlah minimum support = 100%, dengan rumus sebagai berikut[11][12]:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} * \quad (1)$$

Pembentukan pola frekuensi dua itemsets, dibentuk dari items-items jenis yang memenuhi support minimal yaitu dengan cara mengkombinasi semua item kedalam pola dua kombinasi, hasil pembentukan pola kombinasi 2 itemset yang dibentuk dari tabel diatas dan Proses Pembentukan C2 atau disebut dengan 2 itemset, dapat diselesaikan dengan rumus berikut:

$$\text{Support}(A,B) = P(A \cap B)$$

$$\text{Support (A B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}} \quad (2)$$

3. HASIL DAN PEMBAHASAN

Analisis data dilakukan dengan mengumpulkan data yang dibutuhkan setelah data terkumpul maka dilakukan proses *cleaning* data yang membersihkan data tersebut dari informasi-informasi yang tidak diperlukan oleh sistem, lalu setelah melalui proses *cleaning* data diseleksi yang bertujuan mentransformasikan data ke suatu format yang membuat proses pengerjaan menjadi lebih mudah dan efektif dengan tujuan agar mendapatkan hasil yang akurat. Prosedur untuk melakukan *Apriori* [5] Sebagai Berikut :

1. Lakukan penarikan contoh acak berukuran n dengan pemulihan pada gugus data. Tahap ini adalah tahapan *bootstrap*.
2. Dengan menggunakan contoh *bootstrap*, pohon dibangun sampai mencapai ukuran maksimum (tanpa pemangkasan). Pembangunan pohon dilakukan dengan menerapkan *random feature selection* pada setiap proses pemilihan pemilah, yaitu m peubah penjelas dipilih secara acak dimana $m \ll p$, lalu pemilah terbaik dipilih berdasarkan m peubah penjelas tersebut.
3. Ulangi langkah 1 dan 2 sebanyak k kali, sehingga terbentuk sebuah hutan yang terdiri atas k pohon.

3.1 Analisa Pencarian Pola Frekuensi

Data persediaan bahan kimia yang digunakan dalam penelitian ini ialah data persediaan yang ada pada tahun 2017 sampai dengan 2018, data diperoleh dari Laboratorium Forensik Cabang Medan yang tersedia. Dalam penelitian ini data persediaan bahan kimia yang akan diolah pada sistem data mining dengan menggunakan metode *Apriori*, terdapat 49 jenis bahan kimia yang digunakan di laboratorium forensic Medan. Adapun data bahan kimia yang akan di proses pada penelitian ini dapat dilihat pada tabel 1 dimana data persediaan bahan kimia pada Laboratorium Forensik Cabang Medan seperti berikut ini:

Tabel 1. Data Persedian Bahan Kimia Tahun 2017

No	Bahan Kimia Cair	Satuan	Masuk	Keluar	Total
1	2- Propanol for analysis	Botol (2,5 ltr)	35	35	35
2	Acetaldehyde	Botol (500 ml)	2	2	2
3	Acetic Acid Glasial for analysis	Botol (2,5 ltr)	1	1	1
4	Ammonia for analysis	Botol (2,5 ltr)	6	6	6
5	Ammonium Chlorida for analysis	Botol (1 Kg)	1	1	1
6	Chloroform for analysis	Botol (2,5 ltr)	70	70	70
7	Dichloromethane for Analysis	Botol (2,5 ltr)	1	1	1
8	Ethanol absolute for analysis	Botol (2.5 ltr)	4	4	4
9	Ethyl Acetate chromatography	Botol (1 ltr)	6	6	6
10	Formaldehyde solution min 37% GR for analysis	Botol (2,5 ltr)	5	5	5
11	Galat Acid	Botol (25 gr)	1	1	1
12	Hydrogen Peroxide 30% for analysis	Botol (1 ltr)	1	1	1
13	Methanol for analysis	Botol (2,5 ltr)	12	12	12
14	N- Hexan for analysis	Botol (2,5 ltr)	6	6	6
15	Nitric acid 65 % for analysis	Botol (2,5 ltr)	1	1	1
16	Petroleum Benzine for analysis	Botol (1 ltr)	1	1	1
17	Petroleum ether for denaturation	Botol (1 ltr)	2	2	2
18	Pyridine for analysis	Botol (500 ml)	1	1	1
19	Selenius Acid	Botol (10 gr)	1	1	1
20	1-Naphthyl Phosphate Disodium Salt	Botol (1 gr)	2	2	2
21	Acetonitrile for analysis	Botol (2.5 L)	1	1	1
22	Alizarin	Botol (100 gr)	1	1	1
23	Asam Pikrat	Botol (100 gr)	1	1	1
24	Barium Chlorida	Botol (500 gr)	1	1	1
25	Cobalt Tyocyanate	Botol (5 gr)	2	2	2
26	Copper Nitrate Trihydrate for	Botol (250 gr)	1	1	1
27	Copper sulfate Pentahydrate	Botol (250 gr)	1	1	1
28	Dhypenylamine	Botol (5 gr)	1	1	1
29	5-(4nitro-phenyl)-2,4-Diphenyl-2,4-dihydro(1,2,4)triazole-3-thione (C20H14N4O2S)	1 Ea / 25 mg	1	1	1
30	Fast Blue BB Salt	Botol (100 gr)	5	5	5
31	Ferro Cyanida	Botol (1 kg)	1	1	1
32	Iron (III) Chloride Hexahydrate for analysis	Botol (250 gr)	2	2	2
33	Leuco Malachite Green	Botol (25 gr)	1	1	1
34	Methylen Blue	Botol (10 gr)	2	2	2
35	Naphthylethylenediamine	Botol (5 gr)	1	1	1
36	Naphtol for syntesis	Botol (1 Kg)	1	1	1
37	Nitron	Botol (5 gr)	1	1	1
38	Potassium Hidroxide (pellets pure)	Botol (1 kg)	1	1	1
39	Potassium Iodida for analysis	Botol (1 kg)	1	1	1
40	Serbuk Zn for anaylis	Botol (500 gr)	1	1	1
41	Silver Nitrate for analysis	Botol (100 gr)	1	1	1
42	Sodium Carbonate	Botol (500 gr)	2	2	2
43	Sodium Chloride	Botol (1 kg)	2	2	2
44	Sodium Dithionite for analysis	Botol (500 gr)	1	1	1

No	Bahan Kimia Cair	Satuan	Masuk	Keluar	Total
45	Sodium Dodecyl Sulfate	Botol (1 kg)	1	1	1
46	Sodium Hidroksida pellet for analysis	Botol (1 kg)	2	2	2
47	Sodium Molibdat dehydrate for analysis	Botol (250 gr)	1	1	1
48	Sulfanilat	Botol (100 gr)	1	1	1
49	Zink Acetate	Botol (250 gr)	1	1	1

Berikut ini adalah penyelesaian dengan kasus berdasarkan data yang sudah disediakan pada tabel 1. Proses pembentukan C1 atau disebut dengan 1 itemset dengan jumlah minimum support = 100%, dengan rumus sebagai berikut:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A} *}{\text{Total Transaksi}}$$
Tabel 2. Nilai Koefisien

No	Nama Barang	Perhitungan	Support
1	2- Propanol for analysis	35/198*100%	17,6%
2	Acetaldehyde	2/198*100%	1%
3	Acetic Acid Glasial for analysis	1/198*100%	0,5%
4	Ammonia for analysis	6/198*100%	3%
5	Ammonium Chlorida for analysis	1/198*100%	0,5%
6	Chloroform for analysis	70/198*100%	35%
7	Dichloromethane for Analysis	1/198*100%	0,5%
8	Ethanol absolute for analysis	4/198*100%	2%
9	Ethyl Acetate chromatography	6/198*100%	3%
10	Formaldehyde solution min 37% GR for analysis	5/198*100%	2,4%
11	Galat Acid	1/198*100%	0,5%
12	Hydrogen Peroxide 30% for analysis	1/198*100%	0,5%
13	Methanol for analysis	12/198*100%	6,06%
14	N- Hexan for analysis	6/198*100%	3%
15	Nitric acid 65 % for analysis	1/198*100%	0,5%
16	Petroleum Benzine for analysis	1/198*100%	0,5%
17	Petroleum ether for denaturation	2/198*100%	1%
18	Pyridine for analysis	1/198*100%	0,5%
19	Selenius Acid	1/198*100%	0,5%
20	1-Naphtyl Phosphate Disodium Salt	1/198*100%	0,5%
21	Acetonitrile for analysis	1/198*100%	0,5%
22	Alizarin	1/198*100%	0,5%
23	Asam Pikrat	1/198*100%	0,5%
24	Barium Chlorida	1/198*100%	0,5%

2. Kombinasi Tingkat Koefisien

$$\text{Support (A B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}}$$

Tabel 3. Kombinasi Tingkat Koefiseien

No	Nama Barang	Perhitungan	Support
1	Petroleum ether for denaturation	2/198*100%	1%
2	Pyridine for analysis	1/198*100%	0,5%
3	Selenius Acid	1/198*100%	0,5%
4	1-Naphtyl Phosphate Disodium Salt	1/198*100%	0,5%
5	Acetonitrile for analysis	1/198*100%	0,5%
6	Alizarin	1/198*100%	0,5%
7	Asam Pikrat	1/198*100%	0,5%
8	Barium Chlorida	1/198*100%	0,5%

3 Pembentukan Aturan Assosiation Rule

$$\text{Confidence} = p(A|B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi Mengandung A}}$$

Tabel 4. Aturan metode Apriori yang Memenuhi Confidence Minimal

No	Nama Barang	Confidence %
6	Formaldehyde solution min 37% GR for analysis	5 2,4%

No	Nama Barang	Confidence %
7	Methanol for analysis	12 6,06%
8	N- Hexan for analysis	6 3%

Hasil dari metode Apriori menunjukkan nilai dalam bentuk presentase dimana bahan kimia yang sering digunakan di laboratorium forensic medan, hasil yang ditampilkan mulai dari 2% hingga 35 %, yang artinya penggunaan bahan kimia yang sering cepat habis yaitu *Chloroform for analysis* dengan nilai presenntasi 35 % disusul dengan penggunaan bahan kimia lainnya seperti 2- *Propanol for analysis* dengan nilai presentasi yang digunakan 17,6%.

4. KESIMPULAN

Berdasarkan uraian yang telah dibahas disimpulkan data yang digunakan dalam pengujian penentuan persediaan barang bahan kimia di Laboratorium Forensik Cabang Medan berdasarkan data bahan kimia yang sering digunakan dan yang diperlukan selama 1 (satu) tahun sebelumnya, sehingga data tersebut diolah dan menghasilkan sistem yang baru demi meningkatkan pelayanan persediaan barang di Laboratorium Forensik Cabang Medan. Algoritma *Apriori* sangat baik digunakan untuk menentukan *stock* bahan kimia berdasarkan tata letak barang di Laboratorium Forensik Cabang Medan sehingga mempermudah bagian gudang atau *stock* dalam melakukan pengecekan *stock* minimum order dan maksimum order jika terjadi kekurangan atau pengurangan bahan kimia setiap bulannya.

REFERENCES

- [1] R. M. Simanjorang, “Implementation of Apriori Algorithm in Determining the Level of Printing Needs,” *Infokum*, vol. 8, no. 2, Juni, pp. 43–48, 2020.
- [2] H. Widayu, S. Darma, N. Silalahi, and Mesran, “Data Mining Untuk Memprediksi Jenis Transaksi Nasabah Pada Koperasi Simpan Pinjam Dengan Algoritma C4.5,” *Issn 2548-8368*, vol. Vol 1, No, no. June, p. 7, 2017.
- [3] Moh.Sholik and A. Salam, “Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-commerce OrderMas,” *Techno.COM*, vol. 17, no. 2, pp. 158–170, 2018.
- [4] R. A. Saputra, “Komparasi Algoritma Klasifikasi Data Mining Untuk Memprediksi Penyakit Tuberculosis (Tb): Studi Kasus Puskesmas Karawang,” *Semin. Nas. Inov. dan Tren*, no. April, pp. 1–8, 2014.
- [5] N. R. Yunus and A. Rezki, “Kebijakan Pemberlakuan Lock Down Sebagai Antisipasi Penyebaran Corona Virus Covid-19,” *SALAM J. Sos. dan Budaya Syar-i*, vol. 7, no. 3, 2020.
- [6] D. Evanko, “Optical imaging of the native brain,” *Nat. Methods*, vol. 7, no. 1, p. 34, 2010.
- [7] E. Buulolo, *Data Mining Untuk Perguruan Tinggi*, 1st ed. Yogyakarta: Deepublish, 2020.
- [8] A. Nursikuwagus and T. Hartono, “Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis Web,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 7, no. 2, p. 701, 2016.
- [9] Y. Yendrizal, “Data Mining Penjualan Tanaman Hias dengan Algoritma APRIORI Pada Toko Flores Elishabet,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 472, 2020.
- [10] D. P. Utomo and M. Mesran, “Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung,” *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020.
- [11] R. Rismanto, L. Darmawan, and A. Prasetyo, “Penerapan Algoritma Apriori Dalam,” vol. 04, no. 02, pp. 97–102, 2017.
- [12] M. Yetri and S. Yakun, “Data Mining Untuk Analisis Pola Pemilihan Menu Pada Penang Corner Cafe Dan Resto Menggunakan Algoritma Apriori,” *J-SISKO TECH J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 1, no. 2, pp. 114–123, 2018.