**MANAJEMEN TEKNOLOGI INFORMASI**

**ANALISIS JURNAL IMPLEMENTASI KNOWLEDGE IN DATABASES (KDD) PADA APOTEK ATAU RUMAH SAKIT**

****

**Oleh :**

**AINUN MARDLIYYAH**

**09031281621046**

**PROGRAM STUDI SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS SRIWIJAYA**

**2018**

**Knowledge Discovery in Databases (KDD)**

Knowledge Discovery and Data Mining (KDD) adalah proses yang dibantu oleh komputer untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan yang berguna. Data mining tools memperkirakan perilaku dan tren masa depan, memungkinkan bisnis untuk membuat keputusan yang proaktif dan berdasarkan pengetahuan. Data mining tools mampu menjawab permasalahan bisnis yang secara tradisional terlalu lama untuk diselesaikan. Data mining tools menjelajah database untuk mencari pola tersembunyi, menemukan infomasi yang prediktif yang mungkin dilewatkan para pakar karena berada di luar ekspektasi mereka.

Pada proses *Knowledge Discovery Database (KDD)* di atas terdapat beberapa tahapan yaitu sebagai berikut:

1. Seleksi Data *(Data Selection)*

*Selection* (seleksi/pemilihan) data merupakan sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery Database (KDD)* dimulai. Data seleksi yang akan digunakan untuk proses *Data Mining ,* disimpan dalam suatu berkas, terpisah dari basis data operasional.

1. Pemilihan Data *(Preprocessing / Cleaning)*

Proses Preprocessing mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data, seperti kesalahan cetak *(tipografi).* Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *KDD,* seperti data atau informasi *eksternal*.

1. Transformasi *(Transformation)*

Pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang *valid* atau siap untuk dilakukan proses *Data Mining.*

1. *Data Mining*

Pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.

1. Interpretasi/Evaluasi *(Interpratation/Evaluation)*

Pada fase terakhir ini yang dilakukan adalah proses pembentukan keluaran yang bersumber pada proses *Data Mining* Pola informasi (Nofriansyah, 2014).

**ANALISIS JURNAL IMPLEMENTASI KNOWLEDGE IN DATABASES (KDD) PADA APOTEK ATAU RUMAH SAKIT**

1. **Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan**

Dalam persaingan di dunia bisnis, khususnya industri Apotek, menuntut para pengembang untuk menemukan suatu strategi jitu yang dapat meningkatkan penjualan barang. Salah satu cara mengatasinya adalah dengan tetap tersediaannya berbagai jenis alat-alat kesehatan secara kontinu digudang Apotik. Untuk mengetahui alat-alat kesehatan apa saja yang dibeli oleh para konsumen, dilakukan teknik analisis keranjang pasar yaitu analisis dari kebiasaan membeli konsumen. Penerapan Algoritma Apriori, membantu dalam membentuk kandidat kombinasi item yang mungkin, kemudian dilakukan pengujian apakah kombinasi tersebut memenuhi parameter support dan conﬁdence minimum yang merupakan nilai ambang yang diberikan oleh pengguna.

Di Apotek Kelambir – 2 Medan aktivitas transaksi dan pelayanan terhadap konsumen setiap harinya semakin lama semakin tinggi, sehingga tanpa disadari hal ini dapat menimbulkan tumpukan data yang semakin besar Sistem informasi persediaan alat-alat kesehatan di Apotek Kelambir – 2 Medan sudah terkomputerisasi artinya semua yang berhubungan dengan aktifitas penjualan dan pembelian menggunakan komputer yang berbasis jaringan dengan database terpusat kedalam satu server. Sistem informasi persediaan barang di Apotek Kelambir – 2 Medan terdiri atas bagi-bagian data obat dan alat-alat kesehatan, gudang farmasi, pelayanan resep obat dan penjualan. Ditampilan gudang farmasi terdiri atas bagian persediaan, laporan pembelian, laporan penjualan sedangkan data alat-alat kesehatan keluar terdiri atas attribute nomor, tanggal, nomor slip, nama, jumlah barang, harga, dan total harga dan setiap transaksi penjualan minimal terdiri atas satu jenis alat-alat kesehatan. Untuk mendapat analisa data maka data penjualan di export kedalam database Microsoft excel karena database Microsoft excel bersifat spreadsheet sehingga sangat mendukung dalam analisa data. Microsoft excel dalam analisis data digunakan sebagai database karena Microsoft excel sangat mendukung beberapa aplikasi data mining, dimana aplikasi (perangkat lunak) data mining digunakan sebagai tempat uji coba atau implementasi. Berhubungan dengan studi kasus pada Apotik Kelambir – 2 Medan, dapat dilakukan analisis diantaranya analisis data, Analisa Pola Frekuensi Tinggi, Pembentukan Pola Kombinasi dua items, dan Pembentukan Pola Aturan Asosiasi. Tahap dalam menganalisa data dengan algoritma Apriori pada penjualan (data alat-alat kesehatan) dimulai dengan menyeleksi dan membersihkan data data yang akan dianalisis, kemudian mencari semua jenis item nama alat-alat kesehatan yang ada didalam list transaksi penjualan, selanjutnya mencari jumlah setiap item yang ada pada setiap transaksi penjualan (alatalat kesehatan). Sesuai dengan support yang telah ditentukan maka terseleksilah beberapa item data yang memenuhi minimal support sesuai dengan jumlah item barang didalam transaksi, itu disebut pembentukan kombinasi satu item. Tahap selanjutnya dari item-item alat-alat kesehatan yang telah terseleksi dibentuk lah kombinasi dua item, maka terbentuk lah beberapa item data dengan kombinasi 2 item yang berbeda, dengan support yang ditentukan maka terseleksi lah beberapa data dua item, ini disebut pembentukan kombinasi dua item. Demikian seterusnya sampai kombinasi batas maksimal item transaksi.

Untuk membuktikan data-data yang telah dihasilkan berupa pola hubungan kombinasi antar items dan rules-rules asosiasi sesuai dengan Algoritma Apriori maka perlu dilakukan pengujian dengan menggunakan suatu aplikasi. Pengujian pertama yang dilakukan dengan aplikasi Tanagra adalah dengan menghitung frekuensi items atau frequent itemsets. Pengujian yang dilakukan pertama adalah pengujian untuk menghasilkan jenis item dan pola kombinasi dua items. Pengujian yang dilakukan selanjutnya adalah pengujian untuk menghasilkan rules-rules, dimana rules-rules tersebut terbentuk dari pola kombinasi items pada pengujian sebelumnya.

Data Mining dapat di implementasikan dengan menggunakan database penjualan alat-alat kesehatan karena dapat menemukan kecenderungan pola kombinasi itemsets sehingga dapat dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok jenis barang apa yang diperlukan kemudian. Penerapan Algoritma Apriori pada teknik Data Mining sangat efisien dan dapat mempercepat proses pembentukan kecenderungan pola kombinasi itemset hasil penjualan alat-alat kesehatan di Apotek Kelambir-2 Medan, yaitu dengan support dan confidence tertinggi adalah Stick Asam Urat - Stick Gula dan Stick Colestrol- Stick Gula.

1. **Implementasi Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus : Apotik Rumah Sakit Estomihi Medan)**

Proses pembentukan pola kombinasi itemsets dan pembuatan rules dimulai dari analisis data. Data yang digunakan adalah data transaksi penjualan obat, kemudian dilanjutkan dengan pembentukan pola kombinasi itemsets dan dari pola kombinasi itemsets yang menarik terbentuk association rules. Dengan studi kasus pada apotik rumah sakit Estomihi Medan, dapat dilakukan analisa terhadap data khusus data penjualan (data obat keluar) dengan salah satu tujuan adalah untuk menemukan pola kombinasi penjualan obat dan hubungan antar item jenis obat didalam transaksi.

Tahap dalam menganalisa data dengan algoritma apriori pada penjualan (data obat keluar) dimulai dengan menyeleksi dan membersihkan data-data yang akan dianalisis, kemudian dicari semua jenis item nama obat yang ada didalam list transaksi penjualan, selanjut dicari jumlah setiap item yang ada pada semua data transaksi penjualan (data obat keluar). Pembentukan pola kombinasi didasarkan pada nilai support minimal, Jika nilai support minimal terpenuhi dan pola kombinasi itemsets yang ada lebih dari pada satu pola kombinasi, maka pola kombinasi itemsets yang berikutnya bisa bentuk. Setelah selesai pembuatan pola maka langkah selanjutnya pembentuk rules association, rules yang akan dihasilkan dibentuk dari pola kombinasi itemsets yang memenuhi support minimal.

Pembentukan pola frekuensi dua itemsets, dibentuk dari items-items jenis obat yang memenuhi support minimal yaitu dengan cara mengkombinasi semua items kedalam pola dua kombinasi. Setelah semua pola frekuensi tinggi ditemukan, baru dicari association rules yang memenuhi syarat minimum confidence, dengan menghitung confidence aturan asosiati A ke B. Pembentukan association rules adalah menganalisis pola frekuensi tinggi, tahap ini mencari kombinasi yang memenuhi syarat minimum dari support dalam database.

Setelah melakukan pengujian untuk menghasilkan jenis items dan pola kombinasi, mulai dari kombinasi dua items sampai dengan pola kombinasi lima items, dimana rules-rules tersebut terbentuk dari pola kombinasi items pada pengujian sebelumnya. Pengujian yang dilakukan adalah pengujian untuk menghasilkan rules-rules yang terbentuk dari pola kombinasi dua items.

Dengan algoritma apriori dan pengujian dengan aplikasi Tanagra menghasilkan pola kombinasi itemsets dan rules sebagai ilmu pengetahuan dan informasi penting dari data penjualan (data obat keluar). Teknik data mining denga algoritma apriori dapat diimplementasikan pada sistem persediaan dengan data yang digunakan adalah data penjualan (data obat keluar). Hasil yang diperoleh dengan algoritma apriori dan pengujian dengan aplikasi adalah salah satunya, Pola kombinasi yang paling tinggi supportnya adalah pola kombinasi Ketorolac 3% Inj 30 mg/ml, Ringer Lactate Larutan, sedangkan pola kombinasi yang paling banyak itemnya adalah kombinasi itemsets Ceftriaxone Inj 1.0g, Ketorolac 3% Inj 30 mg/ml, Metronidazole Inj 5mg/ml - 100ml, Ringer Lactate Larutan dan Pitogin Inj 10 IU/ ml /sintosin. Data mining dengan algoritma apriori memiliki kelemahan karena harus melakukan scan database setiap kali iterasi, sehingga waktu bertambah setiap kali iterasi. 5. Banyaknya asosiasi antar data dan pola kombisi dan rules yang makin akurat, didapat berdasarkan volume data dan level confidence dan support yang bervariasi.

1. **Aplikasi Data Mining dengan Menggunakan Teknik ARM untuk Pengolahan Informasi Rendemen Obat**

Pada penelitian ini akan dibangun aplikasi randemen yang berasal dari data mining yang telah terbentuk dan membantu perusahaan dalam mengelompokan data dan mengolah data tersebut menjadi suatu informasi yang dibutuhkan oleh badan POM sebagai pengawas obat. Sedangkan teknik mining yang digunakan pada penelitian ini untuk menemukan aturan assosiatif antara suatu kombinasi item adalah association rule mining (ARM). Kombinasi item yang akan dijadikan sebagai index keterkaitan antara database adalah frequent itemset Dengan adanya ferquent itemset memudahkan penentuan atribut yang akan dimunculkan dalam menampilkan informasi rendemen.

Pada penelitian menerapkan linier sequential, yang terdiri dari pengumpulan data dan membentuk data mining, kemudian tahap analisa dan perancangan menggunakan teknik association rule untuk memudahkan penentuan keterkaitan data dan kebutuhan data yang akan ditampilkan, dan tahap akhir pengkodean untuk membuat aplikasi rendemen.

Laporan akhir yang diharapkan dari aplikasi ini adalah berupa laporan obat bulanan, triwulan, semester dan tahunan berdasarkan kategori homogen (satu jenis obat) dan heterogen (campuran yaitu syrup, larutan, kapsul dan injeksi). Hasil akhir dari aplikasi ini diperoleh informasi yang dibutuhkan oleh pimpinan untuk menentukan kebijakan optimal dalam menentukan efisiensi dan efektivitas waktu. Serta nilai keuntungan yang dicapai dalam setiap produksi untuk masing-asing obat. Aplikasi inipun membantu memudahkan pebuatan laporan yang terkomputerisasi dan optimal dalam penyususnan laporan.

1. **Analisis Pola Asosiasi Dan Sekuensial Data Rekam Medis Rsud Dr. H. Slamet Martodirdjo Pamekasan Dengan Teknik Data Mining Menggunakan Algoritma Apriori**

Data rekam medis yang akan diolah adalah data rekam medis pasien Jamkesmas RSUD Dr.H. Slamet Martodirdjo pada kurun tahun 2010, 2011, dan 2012. Sebelum dilakukan proses data mining, data rekam medis yang diperoleh dari basis data rumah sakit dilakukan tahap preprocessing. Mekanisme preprocessing pada tahap integration dan transformation dilakukan pengumpulan data dari tabel-tabel yang ada pada aplikasi resmi di rumah sakit menjadi tabel induk. Berdasarkan data selection dalam preprocessing maka atribut-atribut yang dipakai untuk melakukan data mining adalah: PASIEN\_ID, TANGGAL\_PERIKSA, dan DIAGNOSA\_PENYAKIT sebagai input atributnya.

Skenario penelitian dibagi menjadi 3 tahap. Tahap pertama menggunakan dataset rekam medis tahun 2010 untuk dicari pola asosiasi dan sekuensialnya. Tahap kedua dan ketiga masing-masing mengambil dataset rekam medis tahun 2011 dan tahun 2012 untuk juga dicari pola asosiasi dan sekuensialnya. Minimum support count yang digunakan untuk algoritma association rule untuk masing-masing tahun 2010 adalah 8, tahun 2011 adalah 10 dan tahun 2012 adalah 9. Nilai prosentase yang diambil sebagai confidence pada rule sebesar >= 75%. Untuk minimum support count yang digunakan untuk algoritma sequential pattern untuk masing-masing tahun 2010 adalah 10, 2011 adalah 15 dan tahun 2012 adalah 7. Penentuan angka minimum support mempertimbangkan pada jumlah dataset yang diperoleh (populasi) dan efisiensi pada proses generate item.

Setelah dilakukan penelitian terhadap database rekam medis pasien dari tahun 2010 s/d 2012 didapatkan hasil bahwa kasus penanganan penyakit yang terjadi di RS Dr.H. Martodirdjo Pamekasan selama kurun waktu tahun 2010 adalah penyakit yang berkenaan dengan pernafasan baik yang disebabkan oleh bakteri dan non-bakteri, penyakit berkenaan dengan mata yaitu katarak dan kerusakan lensa mata, dan kasus penyakit berkenaan dengan kencing manis. Setelah dilakukan diagnosis dan komentar untuk tiap-tiap kasus penyakit pada rule dan pattern maximal sequence yang dihasilkan pada proses mining sequential pattern, didapatkan bahwa rule dan pattern tersebut masuk akal dan memang merupakan hubungan sebab akibat dari terjadinya suatu penyakit. Hal tersebut dikarenakan diagnosis pada status pasien hanya dokter yang berhak memberikan pengkodean. Adapun mekanisme pemeriksaan pasien oleh dokter adalah pertama dilakukan proses anamnesi, yaitu proses tanya jawab dengan dokter untuk mengetahui keluhan pasien. Setelah itu dilakukan pemeriksaan fisik kemudian dokter dapat menentukan diagnosis terhadap pasien. Setelah diagnosis diketahui, langkah selanjutnya dokter menentukan memberi tindakan atau cukup memberi resep obat. Komentar dari dokter menguatkan bahwa hasil dari percobaan tidak bertentangan dengan realitas pada kasus penyakit yang menjadi hubungan sebab akibat yang terjadi di lokasi penelitian. Keterangan dari dokter yang bertanggungjawab pada Unit Gawat Darurat (UGD) menyatakan bahwa pattern yang dihasilkan memberikan gambaran secara umum kasus penyakit yang banyak ditangani di RSUD Dr.H.Slamet Martodirdjo Pamekasan.

1. **Penggunaan Teknik Data Mining untuk Manajemen Resiko Sistem Informasi Rumah Sakit**

Sistem informasi kesehatan untuk rumah sakit mulai disimpan secara elektronik sebagai sistem informasi rumah sakit dan meningkat secara tajam. Data yang disimpan pada sistem informasi ini adalah seluruh informasi yang menunjang proses bisnis dari sebuah rumah sakit, mulai dari data laboratorium sampai data pasien. Data yang disimpan akan digunakan kembali untuk kebutuhan-kebutuhan tertentu, oleh karena itu dibutuhkan sebuah teknik untuk penggunaan data kembali, salah satunya dengan penggalian data atau data mining.

Pendekatan teknik penggalian data dalam hal ini digunakan untuk memanfaatkan data resiko yang diekstraksi dari sistem informasi. Teknik penggalian resiko (risk mining) dikenal sebagai pendekatan teknik penggalian data untuk menangani masalah. Teknik ini terdiri dari tiga buah proses, yaitu pendeteksian resiko (risk detection), klarifikasi resiko (risk clarification), dan pemanfaatan resiko (risk utilization).

Pendeteksian resiko (risk detection), pada proses ini dilakukan penggalian pola atau tipe lain dari informasi yang tidak terduga untuk domain expert. Hal ini dilakukan karena pola atau tipe lain informasi yang tidak terduga untuk domain expert akan sangat penting untuk mendeteksi kemungkinan dari kecelakaan yang berskala besar.

Klarifikasi resiko (risk clarification), pada proses ini, fokus ditujukan pada detected risk information yang dihasilkan dari proses risk detection. Domain expert atau penggali data dapat berfokus pada klarifikasi dari pemodelan mekanisme yang tersembunyi dari resiko. Apabila domain expert membutuhkan informasi dengan granulasi yang lebih baik, dapat dilakukan pengumpulan data yang lebih banyak dengan informasi yang rinci, dan mengaplikasikan teknik penggalian data kembali terhadap data yang baru dikumpulkan.

Pemanfaatan resiko (risk utilization), pada tahap ini, dilakukan validasi terhadap model resiko yang dihasilkan dari proses Risk Clarification. Kita harus mengevaluasi clarified risk information di lingkungan nyata untuk mencegah resiko terjadi. Jika informasi resiko tersebut tidak cukup untuk melakukan pencegahan, maka dibutuhkan analisis yang lebih mendalam.

Pada bagian analisis, telah dilakukan analisis perbandingan antara pendekatan teknik penggalian data dan metode manajemen resiko yang diperkenalkan oleh NIST. Berdasarkan analisis tersebut, diketahui bahwa penggunaan pendekatan teknik penggalian data untuk manajemen resiko dapat menghasilkan pemodelan keadaan sistem informasi yang lebih akurat berdasarkan rules yang telah ditentukan sebelumnya. Selain itu, proses pemodelan ini juga dapat dilakukan dengan efektif dan efisien. Hasil pemodelan yang dihasilkan oleh risk mining kemudian digunakan untuk menentukan tindakan untuk mengurangi resiko yang mungkin terjadi. Sementara itu, proses manajemen resiko yang diperkenalkan oleh NIST cenderung lebih lengkap dan rinci. Dengan demikian, kedua pendekatan ini dapat digabungkan untuk memberikan hasil analisis yang lebih baik dan lebih cepat. Misalnya, tahapantahapan yang digunakan dalam manajemen resiko mengikuti tahapan-tahapan pada manajemen resiko yang diperkenalkan oleh NIST, sementara teknik yang digunakan dalam setiap tahapnya dilakukan dengan pendekatan teknik penggalian data. Dengan demikian, hasil analisis yang didapatkan akan lebih akurat dan cepat.