**IMPLEMENTASI ALGORITMA APRIORI DALAM PREDIKSI DATA MINING STOK BUKU**

**Antoni1, Solly Aryza2, Abdul Razak Nasution3, J. E. Melky Purba4**

**1Universitas Islam Sumatera Utara, Medan, Sumatera Utara, Indonesia**

**2Fakultas Sains dan Tekhnologi, Prodi Teknik Elektro, Universitas pembangunan Panca budi**

**3,4 Fakultas Sosial dan Sains, Universitas Pembangunan Pancabudi**

**ABSTRAK**

Penerapan Data Mining algoritma Apriori dalam memprediksi stok buku dapat digunakan untuk membantu perpustakaan dalam memprediksi jenis buku apa saja yang sering dipinjam atau tidak sering dipinjam pada perpustakaan. Sehingga dapat diperoleh informasi jenis buku apa saja yang harus memiliki stok yang banyak, sedikit dan lain sebagainya. Dengan memanfaatkan perangkat lunak yang dirancang untuk melakukan Strategi yang dapat menentukan pola persediaan stok buku berdasarkan data peminjaman merupakan salah satu cara yang tepat untuk mengetahui minat peminjam. Penggunaan algoritma apriori merupakan salah satu cara yang tepat untuk mengetahui hal tersebut. Dengan mengetahui minat peminjam, maka menegement bisa menentukan pola persediaan stok buku sesuai dengan permintaan peminjam. Hasil dari penelitian ini adalah, perpustakaan dapat menyediakan buku yang lebih banyak dipinjam berdasarkan stok buku yang ada. Sehingga peminjam akan meningkat bagi perpustakaan dikarenakan peminjam merasa puas dengan stok yang telah disediakan oleh perpustakaan.

***Kata Kunci:*** *Data Mining, Metode Apriori, Prediksi Stok Buku.*

**PENDAHULUAN**

Data mining menggunakan beberapa teknik dalam mengambil keputusan. Teknik data mining yang dapat dipergunakan adalah untuk memprediksi stok buku perpustakaan berdasarkan permintaan peminjam masih mengalami kendala, dikarenkan sistem yang kurang mendukung. Dibutuhkan sistem informasi yang dapat membantu perpustakaan untuk mengetahui buku seperti apa yang harus ditingkatkan sesuai dengan permintaan pelanggan.

Penerapan yang sesuai untuk mengatasi masalah tersebut adalah dikembangkannya konsep data mining. Algoritma apriori adalah algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *Boolean*. (Sensuse, 012). *Algoritma Apriori* adalah algoritma paling terkenal untuk menemukan pola frekuensi tinggi. Algoritma apriori dapat dengan mudah menentukan barang – barang yang mempunyai kaitan yang lebih dekat hubungannya, sehingga mengetahui bagaimana cara menentukan pola persediaan stok buku berdasarkan kebutuhan permintaan peminjam buku perpustakaan.

Tujuan utama yan ingin dicapai yaitu untuk menentukan persediaan stok buku berdasarkan permintaan peminjam dengan menerapkan metode apriori agar data persediaan stok buku perpustakaan lebih akurat.

**TINJAUAN PUSTAKA**

* 1. **Data Mining**

Data mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Data mining merupakan proses penggalian data dari database yang berukuran besar untuk menentukan informasi penting dan bermanfaat. Klasifikasi adalah salah satu teknik yang ada pada data mining (Sri Wahyuni, 2018).

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai basis data besar (Turban et al, 2005). Data mining merupakan teknik untuk mengeksplorasi data yang banyak sehingga menghasilkan sebuah keputusan yang dapat dijadikan acuan untuk ke depannya.

Data mining adalah sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Data mining juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan (Prasetyo, 2012). Data mining sering juga disebut dengan *Knowledge Discovery in Database* atau disingkat KDD, adalah kegiatan yang meliputi pengumpulan, pemakaian data histori untuk menelusuri data yang ada untuk membangun sebuah model agar dapat mengenali pola data yang lain berukuran besar (Prasetyo, 2012). Berikut adalah tahapan proses KDD :



**Gambar 1.** Tahapan Proses KDD

Enam elemen yang paling esensial dalam teknik pencarian informasi/pengetahuan dalam KDD yaitu (Paresh et al, 2013) :

1. Mengerjakan sejumlah besar data.
2. Diperlukan efesiensi berkaitan dengan volume data.
3. Mengutamakan ketetapan/keakuratan.
4. Membutuhkan pemakaian bahasa tingkat tinggi.
5. Menggunakan beberapa bentuk dari pembelajaran otomatis. Menghasilkan hasil yang menarik.
	1. **Algoritma Apriori**

Algoritma apriori adalah algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *Boolean*. (Sensuse, 012). *Algoritma Apriori* adalah algoritma paling terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola item di dalam suatu database yang memiliki frekuensi atau support di atas ambang batas tertentu tertentu yang disebut dengan istilah minimum support. Pola frekuensi tinggi ini digunakan untuk menyusun aturan hubungan positif dan juga beberapa teknik data mining lainnya.

Algoritma apriori adalah algoritma pengambilan data dengan aturan asosiatif (*Association Rule*) untuk menentukan hubungan asosiatif suatu kombinasi item. Analisis pola frekuensi tinggi dengan algoritma appriori mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam basis data (Hermawati, 2013). Nilai *support* sebuah item diperoleh dengan menggunakan rumus berikut:

$$Support\left(A\right)=\frac{Jumlah Transaksi Mengandung A}{Total Transaksi}$$

Nilai *support* dari 2 item diperoleh dengan menggunakan rumus :

$$Support\left(A,B\right)=P\left(A⋂B\right)$$

$$Support\left(A,B\right)=\frac{\sum\_{}^{}transaksi mengandung A dan B}{\sum\_{}^{}transaksi}$$

Pembentukan aturan asosiasi setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan rumus berikut:

$$Confi=P\left(A\right)=\frac{\sum\_{}^{}transaksi mengandung A dan B}{\sum\_{}^{}transaksi mengandung A}$$

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan *Support* × *Confidence*. Aturan diambil sebanyak n aturan yang memiliki hasil terbesar.

Algoritma apriori adalah algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *Boolean*. (Sensuse, 012). *Algoritma Apriori* adalah algoritma paling terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola item di dalam suatu database yang memiliki frekuensi atau support di atas ambang batas tertentu tertentu yang disebut dengan istilah minimum support. Pola frekuensi tinggi ini digunakan untuk menyusun aturan hubungan positif dan juga beberapa teknik data mining lainnya.

Algoritma apriori adalah algoritma pengambilan data dengan aturan asosiatif (*Association Rule*) untuk menentukan hubungan asosiatif suatu kombinasi item. Analisis pola frekuensi tinggi dengan algoritma appriori mencari kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam basis data (Hermawati, 2013). Nilai *support* sebuah item diperoleh dengan menggunakan rumus berikut:

$$Support\left(A\right)=\frac{Jumlah Transaksi Mengandung A}{Total Transaksi}$$

Nilai *support* dari 2 item diperoleh dengan menggunakan rumus :

$$Support\left(A,B\right)=P\left(A⋂B\right)$$

$$Support\left(A,B\right)=\frac{\sum\_{}^{}transaksi mengandung A dan B}{\sum\_{}^{}transaksi}$$

Pembentukan aturan asosiasi setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan rumus berikut:

$$Confi=P\left(A\right)=\frac{\sum\_{}^{}transaksi mengandung A dan B}{\sum\_{}^{}transaksi mengandung A}$$

Untuk menentukan aturan asosiasi yang akan dipilih maka harus diurutkan berdasarkan *Support* × *Confidence*. Aturan diambil sebanyak n aturan yang memiliki hasil terbesar.

**METODE PENELITIAN**

Dalam metode penelitian ini bertujuan untuk lebih memahami bagaimana mengimplementasikan Data Mining Algoritma apriori dalam memprediksi stok buku perpustakaan.

* 1. **Tahapan Penelitian**

Penelitian akan dilaksanakan berdasarkan tahap penelitian yang ditunjukkan pada Gambar 3.1.

Pengumpulan Data

Analisa Masalah

Identifikasi Masalah

Perumusan Masalah

Implementasi

Pengujian

**Gambar 2.**  Tahapan Penelitian

1. Tahap awal dilakukan pengumpulan data yang bertujuan untuk menyediakan bahan agar proses selanjutya yaitu menganalisa permasalahan yang terjadi pada proses kegiatan pengolahan data buku di perpustakaan.
2. Tahap analisa permasalahan dilakukan untuk dapat melihat seberapa banyak permasalahan yang diperoleh berdasarkan tingkat kepentingan dari permasalahan yang akan diselesaikan.
3. Tahap identifikasi masalah adalah mengelompokan permasalahan yang diperoleh yang sesuai dengan tujuan penelitian.
4. Tahap merumuskan masalah adalah tahap pemilihan masalah yang telah teridentifikasi untuk dapat dirumuskan agar dapat menyelesaikan permasalahan yang sesuai dengan tujuan penelitian.
5. Proses pengolahan data dilakukan terlebih dahulu melakukan identifikasi sering dihadapi oleh pihak perpustakaan, kemudian mendeskripsikan masalah – masalah tersebut untuk diperoleh solusinya. Tahap selanjutnya dilakukan analisa masalah menggunakan teknik data mining dengan algoritma apriori untuk mendapatkan hasil sebagai tujuan yang akan dicapai.
6. Pada tahap ini dilakukan analisis hasil berdasarkan perhitungan algoritma apriori yang dapat digunakan untuk pengambilan keputusan dalam penentuan pola penempatan buku.
7. Tahap akhir adalah menyimpulkan hasil penelitian apakah telah memberikan solusi dari permasalah yang telah diselesaikan.

**HASIL PENELITIAN DAN DISKUSI**

* 1. **Penentuan nilai minimum support**

Untuk menganalisa nilai minimum support dan confidence antara satu item dengan item yang lain. Maka dibutuhkan tahapan-tahapan transaksi yang akan menghasilkan suatu daftar dari satu itemset yang dapat dilihat sebagai berikut :

**Tabel 1.** Daftar 1 *Itemset*

|  |  |  |
| --- | --- | --- |
| No. | Kode | Jlh Transaksi |
|  1. | 684BN10 | 15 |
|  2 | 686BN11  | 18 |
|  3 | 674BN09  | 20 |
|  4 | 675BN12 | 13 |
|  5 | 654BN07 | 26 |
|  6 | 481BU12 | 11 |
|  7 | 473BU12  | 16 |
|  8 | 235BN05 | 25 |
|  9 | 237BN06 | 12 |
| 10 | 597BU13 | 9 |
| 11  | 585BU12  | 22 |
| 12 | 1356BN11 | 14 |
| 13 | 1352BN12 | 21 |

Untuk menghitung nilai support dari setiap jenis buku dapat digunakan rumus sebagai berikut:

***Support (A)* = Jumlah transaksi mengandung A / Total Transaksi) x 100%**

Dari rumus diatas maka dapat dihasilkan nilai support dari masing-masing buku, yaitu:

*Support* (684BN10) = $\frac{15}{29}$ x 100% = 51,72%

*Support* (686BN11) = $\frac{14}{29}$ x 100% = 48,27%

*Support* (674BN09) = $\frac{20}{29}$ x 100% = 68,96%

*Support* (675BN12) = $\frac{13}{29}$ x 100% = 44,82%

*Support* (654BN07) = $\frac{26}{29}$ x 100% = 89,65%

*Support* (481BU12) = $\frac{11}{29}$ x 100% = 37,93%

*Support* (473BU12) = $\frac{16}{29}$ x 100% = 55,17%

*Support* (235BN05) = $\frac{25}{29}$ x 100% = 86,20%

*Support* (237BN06) = $\frac{12}{29}$ x 100% = 41,37%

*Support* (597BU13) = $\frac{9}{29}$ x 100% = 31,03%

*Support* (585BU12) = $\frac{22}{29}$ x 100% = 75,86%

*Support* (1356BN11) = $\frac{14}{29}$ x 100% = 48,27%

*Support* (1352BN12) = $\frac{9}{29}$ x 100% = 31,03%

*Support* (684BN10, 686BN11) = $\frac{15}{29}$ x 100% = 51,72%

*Support* (684BN10, 675BN12) = $\frac{18}{29}$ x 100% = 62,06%

*Support* (684BN10, 235BN05) = $\frac{13}{29}$ x 100% = 44,82%

*Support* (684BN10, 236BN06) = $\frac{11}{29}$ X 100% = 37,93%

*Support* (684BN10, 237BN07) = $\frac{13}{29}$ x 100% = 44,82%

*Support* (684BN10, 597BU13) = $\frac{8}{29}$ x 100% = 27,58%

*Support* (684BN10, 675BN12) = $\frac{6}{29}$ x 100% = 20,68%

*Support* (684BN10, 235BN05) = $\frac{11}{29}$ x 100% = 37,93%

*Support* (684BN10, 237BN06) = $\frac{14}{29}$ x 100% = 48,27%

*Support* (684BN10, 240BN09) = $\frac{9}{29}$ x 100% = 31,03%

*Support* (684BN10, 597BU13) = $\frac{8}{29}$ x 100% = 27,58%

*Support* (675BN12, 235BN05) = $\frac{9}{29}$ x 100% = 31,03%

*Support* (675BN12, 236BN06) = $\frac{15}{29}$ x 100% = 51,72%

*Support* (675BN12, 237BN07) = $\frac{21}{29}$ x 100% = 72,41%

*Support* (675BN12, 597BU13) = $\frac{7}{29}$ x 100% = 24,13%

*Support* (235BN05, 237BN05) = $\frac{12}{29}$ x 100% = 41,37%

*Support* (235BN05, 240BN09) = $\frac{11}{29}$ x 100% = 37,93%

*Support* (235BN05, 597BU13) = $\frac{5}{29}$ x 100% = 17,24%

*Support* (237BN06, 240BN09) =$ \frac{18}{29}$ x 100% = 62, 06%

*Support* (237BN06, 597BU13) = $\frac{7}{29}$ x 100% = 24,13%

*Support* (240BN09, 597BU13) = $\frac{8}{29}$ x 100% = 27,58%

Dari hasil perhitungan support 2 itemset diatas, kemudian itemset mana yang memiliki minimum support diatas 50% untuk kemudian dihitung nilai confidence tiap itemset seperti pada tabel berikut:

**Tabel 2.** Nilai Confidence tiap itemset

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Itemset** | **Jumlah** | **Support** |
| 1 | 684BN10685BN11 | 17 | 58,62% |
| 2 | 684BN10687BN13 | 15 | 51,72% |
| 3 | 685BN11686BN12 | 20 | 68,96% |
| 4 | 685BN11687BN13 | 16 | 55,17% |
| 5 | 685BN11654BN07 | 22 | 75,86% |
| 6 | 685BN111356BN11 | 15 | 51,72% |
| 7 | 686BN12687BN13 | 21 | 72,41% |
| 8 | 687BN13688BN14 | 18 | 62,06% |

* 1. **Penentuan nilai minimum confidence.**

Langkah selanjutnya adalah menghitung *Confidence* dari kandidat 2 item dengan persamaan sebagai berikut :

***Confidence* (A dan B) = (jumlah transaksi mengandung A dan B/ jumlah transaksi mengandung A) x 100%**

*Confidence* (684BN10, 685BN11) = $\frac{17}{19}$ x 100% = 89,47%

*Confidence* (684BN10, 687BN13) = $ \frac{11}{21}$ x 100% = 52,38%

*Confidence* (685BN11, 686BN12) = $ \frac{20}{25}$ x 100% = 80%

*Confidence* (685BN11, 687BN13) = $\frac{16}{21 }$ x 100% = 76,19%

*Confidence* (685BN11, 654BN07) = $\frac{22}{27}$ x 100 % = 81,48%

*Confidence* (685BN11, 1356BN11) = $\frac{15}{27}$ x 100% = 55,55%

*Confidence* (686BN12, 687BN13) = $\frac{21}{25 }$ x 100% = 84%

*Confidence* (687BN13, 688BN14) = $\frac{18}{23}$ x 100% = 78,26%

* 1. **Penentuan Association Rules**

Dari hasil association rules A-> B maka dibuat keterangan untuk setiap rule seperti terlihat pada table sebaai berikut :

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | ***Rule*** | ***Confidence*** | **Keterangan** |
| 1 | {684BN10} - > {685BN11} | 89,47% | 89,47% peminjam buku dengan kode 684BN10, akan meminjam buku dengan kode 685BN11 pada setiap bulannya. |
| 2 | {686BN12} - > {687BN13} | 84% | 84% peminjam buku denan kode 686BN12, akan meminjam buku dengan kode 687BN13 setiap bulannya. |
| 3 | {685BN11} - > {654BN07} | 81,48% | 81,48% peminjam buku dengan kode 685BN11, akan meminjam buku dengan kode 654BN07 pada setiap bulannya. |
| 4 | {685BN11} - > {686BN12} | 80% | 80% peminjam buku dengan kode 685BN11, akan meminjam buku dengan kode 685BN12 pada setiap bulannya. |
| 5 | {687BN13} - > {688BN14} | 78,26% | 78,26% peminjam buku dengan kode 687BN13, akan meminjam buku dengan kode 687BN14 pada setiap bulannya. |
| 6 | {685BN11} - > {687BN13} | 76,19% | 76,19% peminjam buku dengan kode 685BN11, akan meminjam buku dengan kode 867BN13 pada setiap bulannya. |

* 1. **Hasil Penentuan Association Rules**

Dari hasil analisis data diatas dapat disimpulkan bahwa pola peminjaman buku paling banyak terdapat pada buku dengan kode 684BN10 dan 685BN11 dengan persentase peminjaman mencapai 89,47%. Oleh karena itu Perpustakaan Universitas Pembangunan Panca Budi Medan harus mempunyai stok buku dengan kode 68BN10 DAN 685BN11 untuk memenuhi banyaknya peminjam dan mencegah terjadinya kekurangan stok buku dengan kode tersebut yang kan diprediksikan terjadi pada bulan berikutnya.

**KESIMPULAN**

Dari hasil perancangan aplikasi Implementasi data mining dalam memprediksi stok buku perpustakaan menggunakan algoritma apriori berdasarkan permintaan peminjam pada perpustakaan Universitas Pembangunan Panca Budi Medan, maka diperoleh suau kesimpulam sebaai berikut :

1. Untuk menentukan pola persediaan stok buku berdasarkan permintaan peminjam maka diperlukan data peminjam buku pada perpustakaan Universitas Pembangunan Panca Budi Medan
2. Untuk menerapkan metode apriori dalam menentukan pola persediaan stok buku berdasarkan permintaan peminjam maka dilakukan uji coba metode apriori untuk mengetahui pola persediaan stok buku berdasarkan permintaan peminjam pada perpustakaan Universitas Pembangunan Panca Budi Medan
3. Sistem ini dirancang dengan mengunakan metode apriori untuk menentukan pola persediaan stok buku berdasarkan permintaan peminjam pada perpustakaan Universitas Pembangunan Panca Budi Medan
4. Hasil dari pengujian aplikasi yang telah dibangun dapat melihat dan menentukan pola persediaan stok buku berdasarkan permintaan peminjam pada perpustaaan Universitas Pembangunan Panca Budi Medan agar dijadikan sebagai acuan untuk stok buku pada tahun berikutnya.

**DAFTAR PUSTAKA**

Wahyuni, S.2018. *Implementasi Data Mining dalam Memprediksi Stok Barang Mengunakan Algoritma Apriori*. Medan: Jurnal Teknik dan Informatika. Vol. 5, 2089-5490

Wahyuni, S. (2017). *Implementasi Rapidminer Dalam Menganalisa Data Mahasiswa Drop Out*. Medan: Jurnal Ilmiah Abdi Ilmu. Vol. 10, No.2:1899-1902.

Hermawati FA, 2013, *Data Mining*, Andi, Yogyakarta

Hutahaean, Jeferson. 2014. *Konsep Sistem* *Informasi*. Yogyakarta: Deepublish.

Irmansyah, Fariet. 201 3. *Dasar-Dasar* *Algoritma dan Pemrograman*. Yogyakarta: Andi Offset.

Jogiyanto, Hartono M. 2005. *Analisis dan* *Desain Sistem Informasi: Pendekatan* *Terstruktur; Teori dan Praktek Aplikasi* *Bisnis.* Yogyakarta: Andi Offset.

Kristanto, Andri. 2008. *Perancangan Sistem* *Informasi dan Aplikasinya*. Yogyakarta: Geva Media.