

# IMPLEMENTASI DATA MINING DENGAN ALGORITMA APRIORI UNTUK MEMREDIKSI TINGKAT KELULUSAN MAHASISWA

Irham Kurnawan, Fitri Marisa<sup>2</sup>, Dwi Purnomo<sup>3</sup>

wandk89@gmail.com, fitrimarisa@widyagama.ac.id, purnomo@widyagama.ac.id  
Jurusan Teknik Informatika, Universitas Widyagama Malang

---

## Abstract

*Informatics Engineering Program University of Widyagama Malang has many volume of student databases. When properly excavated, so can known patterns or knowledge for a decision-making. Data that can be explored is the understanding of information about graduation students. This research to predict the passing rate of students with more efficient time, which can be known before the student graduated so that can be evaluated in their studies, especially in Informatics Engineering University of Widyagama. In this case using association methods and Apriori algorithm. This method calculates the support value that is the supporting value of an item with big golden rule 60% of the data of course grade. The results of this study to help universities improving the quality of education and help in knowing the information about the graduation rate of students based on the value of the subjects and achievement index obtained by students in Informatics Engineering course University of Widyagama Malang.*

**Keywords:** Data Mining, Apriori, Tingkat Kelulusan

---

## I. PENDAHULUAN

Berdasarkan hasil observasi pada Prodi Teknik Informatika Universitas Widyagama Malang, permasalahan yang terjadi yaitu tidak sedikit mahasiswa yang mengalami kendala dalam masa perkuliahan. Peran dosen dalam mengarahkan mahasiswa yang memiliki beberapa kendala dalam perkuliahan menjadi sangat penting dilakukan. Namun dengan kuantitas mahasiswa yang sangat banyak dan berbagai permasalahan yang berbeda menyebabkan fungsi konsultatif tersebut kurang efektif untuk dilakukan.

Dengan menggunakan algoritma apriori ini bertujuan untuk mengetahui nilai matakuliah yang paling dominan untuk memprediksi tingkat kelulusan mahasiswa. Peneliti berharap dapat membantu dalam mengetahui informasi tentang nilai mata kuliah dan indeks prestasi kumulatif (IPK) yang dapat dijadikan rekomendasi dalam menentukan tingkat kelulusan mahasiswa pada program studi Teknik Informatika Universitas Widyagama Malang. Selain itu, sebagai usaha untuk membantu perguruan tinggi dalam meningkatkan kuantitas mahasiswa pada tahun berikutnya.

## II. TINJAUAN PUSTAKA

### A. Tingkat Kelulusan

Tingkat kelulusan mahasiswa dapat dilihat dari Indeks Prestasi Kumulatif (IPK) yang mencerminkan seluruh nilai yang diperoleh mahasiswa setiap semester yang sedang berjalan. IPK diperoleh dengan menjumlahkan nilai dari semua mata kuliah yang telah diambil dan membaginya dengan total SKS (Satuan Kredit Semester) [1].

### B. Data Mining

*Data mining* adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (*database*) yang sangat besar

sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui [2].

Sedangkan definisi yang lain menerangkan bahwa, *data mining* adalah suatu pencarian dan analisa dari jumlah data yang sangat besar dan bertujuan untuk mencari arti dari pola dan aturan [3].

### C. Association Rule

*Association Rule Mining* merupakan teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item atau untuk menemukan hubungan hal tertentu dalam suatu transaksi data dengan hal lain di dalam transaksi, yang digunakan untuk memprediksi pola [4].

Sedangkan definisi lain menyatakan, *Association Rule Mining* terdiri dari itemset yang sering muncul. *Association Rule Mining* dapat dianalisa lebih lanjut untuk mengungkap aturan korelasi untuk menyampaikan korelasi statistik antara itemsets A dan B [5].

### D. Algoritma Apriori

*Algoritma apriori* adalah algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *Boolean*. Algoritma ini mengontrol berkembangnya kandidat *itemset* dari hasil *frequent itemsets* dengan *support-based pruning* untuk menghilangkan *itemset* yang tidak menarik dengan menetapkan *minsup* [6].

Algoritma apriori juga dapat didefinisikan sebagai suatu proses untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk *support* dan syarat minimum untuk *confidence* [7].

## III. ANALISIS DAN PERANCANGAN

### A. Analisis Sistem

Masalah utama yang terjadi pada Program Studi Teknik Informatika Universitas Widyagama yaitu tidak adanya sarana yang digunakan dalam

mengetahui/memprediksi tingkat kelulusan mahasiswa. Oleh sebab itu, penelitian ini bertujuan untuk mengetahui nilai matakuliah yang paling dominan untuk memprediksi tingkat kelulusan mahasiswa hingga dapat mengurangi jumlah kandidat yang harus dihitung *support*-nya sehingga sistem ini dapat diterapkan dan hasil penelitian ini juga akan dirasakan manfaatnya pada program studi Teknik Informatika Universitas Widyagama.

Dengan demikian, indeks prestasi kumulatif (IPK) yang mencerminkan seluruh nilai yang diperoleh mahasiswa setiap semester yang menjadi tolak ukur besarnya tingkat prestasi yang telah dicapai oleh mahasiswa dapat terus ditingkatkan.

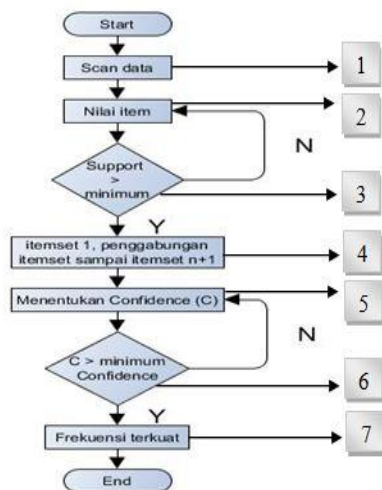
**B. Perancangan Sistem**

Dalam mengatasi permasalahan pada Prodi Teknik Informatika Universitas Widyagama yaitu dengan menggunakan teknik data mining dengan aturan asosiasi algoritma Apriori. Algoritma apriori bertujuan untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk *support* dan syarat minimum untuk *confidence*. Dalam perhitungan algoritma apriori dilakukan perhitungan nilai tiap item data berdasarkan jenis dan nilai dari tiap mata kuliah. Nilai tiap item digunakan dalam menentukan nilai *support* item, dimana jika nilai *support* >= minimum item maka proses perhitungan dapat dilanjutkan untuk menghitung nilai *itemset*.

Nilai *itemset* yang diperoleh dilakukan penggabungan item sehingga diperoleh data *itemset* 1, *itemset* 2, hingga data *itemset* n+1. Kemudian menentukan nilai *confidence*, jika nilai *confidence* yang diperoleh >= *confidence*. Sehingga dari hasil perhitungan dengan menggunakan algoritma apriori tersebut, dapat diperoleh prediksi tingkat kelulusan mahasiswa.

**C. Flowchart Algoritma**

Berikut ini adalah tahapan alur (*flowchart*) perhitungan *Algoritma Apriori* :



**Gambar 1:**

*Flowchart* alur perhitungan *Apriori*

**1. Scan data**

Pada tahap ini dilakukan pengelompokan data mata kuliah. Untuk matakuliah di bedakan menjadi 2 bagian, yaitu matakuliah jenis dan jenis2. Matakuliah jenis merupakan jenis matakuliah sesuai dengan SK Dikti pada program studi teknik informatika yang dibedakan menjadi :

- a) Mata Kuliah Pengembangan Kepribadian (MPK).
- b) Mata kuliah Keilmuan dan Keterampilan (MKK).
- c) Mata kuliah Keahlian Berkarya (MKB).
- d) Mata kuliah Perilaku Berkarya (MPB)
- e) Mata kuliah Berkehidupan Bermasyarakat (MBB).

Mata kuliah jenis2 merupakan kelompok mata kuliah yang lebih spesifik untuk jurusan Informatika, kelompok matakuliah jenis2, yaitu : Algoritma, Dasar IT, *Database*, *Desktop*, *Mobile*, *Pendukung*, *Sistem Informasi*, dan *Web*.

**2. Nilai Item**

Nilai *support* sebuah item dapat kita peroleh dari rumus berikut :

$$Support(A) = \left( \frac{Jumlah\ Transaksi\ mengandung\ A}{Jumlah\ Transaksi} \right) \times 100\% \dots\dots\dots (1)$$

Dari tabel 1, 2 dan 3 berikut, diperoleh pengelompokan *itemset* menjadi 3 (tiga), yaitu :

**Tabel 1 :** Tabel Nilai per item (1-*itemset*)

| Nilai |        |
|-------|--------|
| Item  | Jumlah |
| A     | 23     |
| B+    | 18     |
| B     | 6      |
| C+    | 1      |
| C     | 2      |

**Tabel 2 :** Tabel Matakuliah Jenis per item (1-*itemset*)

| Item | Jumlah |
|------|--------|
| MPK  | 8      |
| MKK  | 10     |
| MKB  | 27     |
| MPB  | 3      |
| MBB  | 2      |

**Tabel 3 :** Tabel Matakuliah Jenis 2 per item (1-*itemset*)

| Item             | Jumlah |
|------------------|--------|
| Database         | 1      |
| Dasar IT         | 4      |
| Pendukung        | 5      |
| Sistem Informasi | 5      |
| Web              | 4      |
| Desktop          | 6      |
| Algoritma        | 9      |
| Mobile           | 1      |

**3. Support Minimum**

Dalam menghitung *support* minimum, yang di gunakan dalam dalam pemilihan minimum frekuensi adalah minimal 60% dari jumlah total maksimal suatu item. Nilai *support* dari 2 item berbeda dapat kita peroleh dari rumus berikut :

$$Support(A, B) = \frac{\sum \text{Transaksi mengandung A dan B}}{\sum \text{Transaksi}} \dots\dots\dots (2)$$

**4. Penggabungan itemset**

Penggabungan itemset dilakukan dari tabel *itemset* matakuliah jenis, matakuliah jenis 2 dan tabel *itemset* nilai, dan mengeluarkan data yang tidak kita butuhkan. Tahap ini memilih *itemset* pertama atau *frequent itemset*. Untuk memilih *frequent itemset* membutuhkan minimum frekuensi kemunculan suatu item hingga diperoleh data berikut:

**Tabel 4 : Tabel per item (4-itemset)**

| Item               | Jumlah |
|--------------------|--------|
| A, MKB, Algoritma  | 3      |
| B+, MKB, Algoritma | 2      |

Dari tabel di atas, dapat kita ketahui bahwa keempat item tersebut lebih sering muncul persamaan dibandingkan item-item yang lainnya. Terdapat item Nilai yang berbeda dari kombinasi 2 item tersebut, dan tidak memungkinkan untuk melanjutkan ke itemset selanjutnya.

**5. Nilai Confidence**

Setelah memilih item-item untuk peluang *support*, selanjutnya kita gunakan untuk mencari peluang *confidence*.

Rumus untuk mencari *confidence* adalah sebagai berikut :

$$Confidence P(B|A) = \frac{\text{Jumlah mengandung A dan B}}{\text{Jumlah mengandung A}} \dots\dots\dots (3)$$

**6. Minimum frekuensi**

Minimum *confidence* yang digunakan dalam pemilihan minimum frekuensi adalah minimal 60% dari nilai *confidence* item. Jika nilai *confidence* > minimum *confidence* maka akan diperoleh frekuensi terkuat.

**7. Final Assosiation Rule**

Setelah diperoleh frekuensi minimum *confidence*, Berikut ini tabel hasil perangkingan nilai berdasarkan jenis matakuliah dan nilai.

**Tabel 5. : Frekuensi Tertinggi**

| ITEM              | C     | Rank |
|-------------------|-------|------|
| Algoritma dan MKB | 77,78 | 1    |

ITEM = penggabungan nilai matakuliah dan jenis matakuliah

C = nilai *confidence* yang diperoleh

Rank = urutan ranking nilai *confidence* yang diperoleh

*Final association rule* menjelaskan tentang support dan confidence dari masing-masing kombinasi 2 itemsets dan 3 itemset. Aturan *final association rule*

dapat terbentuk apabila terbentuknya kombinasi 3 itemset.

Item = penggabungan nilai matakuliah dan jenis matakuliah

*Support* = nilai *support* dari itemset2 dan itemset3

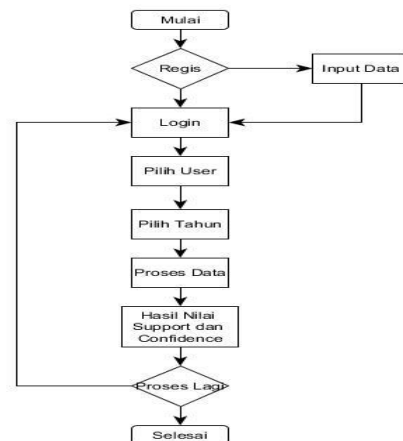
*Confidence* = nilai *confidence* itemset2 dan itemset3

**Tabel 6 : Final Association Rule**

| ITEM           | Support | Confidence | Support x Confidence |
|----------------|---------|------------|----------------------|
| MKB , A        | 110%    | 40,74%     | 44.8%                |
| MKB, B+        | 90%     | 33,33%     | 29,9%                |
| Algoritma, A   | 50%     | 55,55%     | 27,7%                |
| Algoritma, B+  | 20%     | 22,22%     | 4,4%                 |
| Algoritma, MKB | 70%     | 77,78%     | 54,4 %               |

**E. Flowchart Sistem**

*Flowchart* (diagram alir) digunakan untuk mempermudah dalam memahami sistem kerja pada proses. Berikut ini adalah *flowchart* sistem prediksi kelulusan mahasiswa :

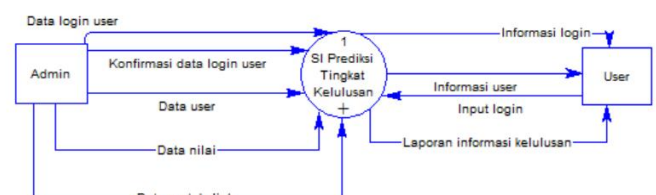


**Gambar 2: Flowchart Sistem**

*Flowchart* tersebut dimulai dengan input data yang dilakukan oleh *user* ketika melakukan login ke sistem. Kemudian data yang di input tersebut akan di proses oleh admin dalam melakukan perhitungan prediksi tingkat kelulusan mahasiswa.

**E. Data Flow Diagram (DFD)**

Berikut merupakan DFD yang menggambarkan bagaimana sistem yang menggunakan metode *asosiasi* algoritma *Apriori* ini berjalan.

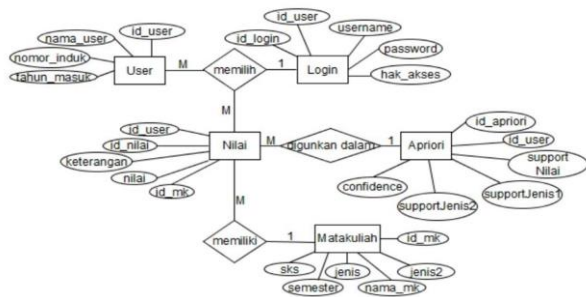


**Gambar 3: DFD Level 0 sistem**

**H. Entity Relationship Diagram (ERD)**

ERD adalah gambar atau diagram yang menunjukkan informasi dibuat, disimpan dan digunakan dalam sistem bisnis. Entitas biasanya menggambarkan

jenis informasi yang sama. Berikut ini adalah ERD yang direkomendasikan pada Sistem Prediksi Kelulusan:

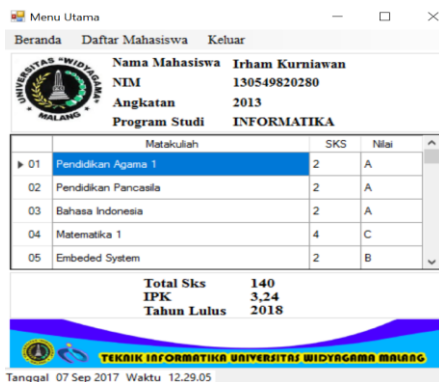


Gambar 4: ERD sistem

IV. HASIL DAN PEMBAHASAN

Berikut ini pembahasan hasil implementasi Sistem data mining menggunakan algoritma apriori untuk memprediksi tingkat kelulusan mahasiswa.

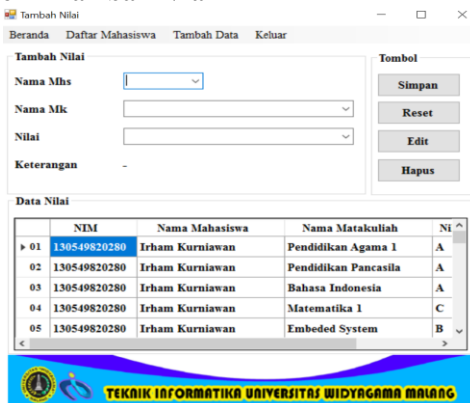
A. Utama



Gambar 5 : Form Utama

Menu utama diatas berisi nama mahasiswa yang melakukan login, nim, angkatan, daftar matauliah, sks, dan nilai yang telah ditempuh, tahun lulus serta nilai IPK mahasiswa tersebut.

B. Form Tambah Nilai



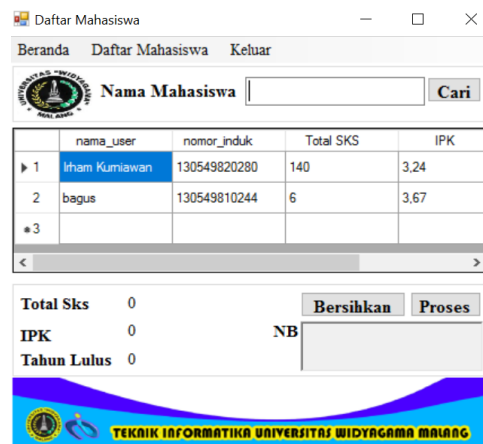
Gambar 6 : Form Tambah Nilai

Menu tambah nilai digunakan untuk menambahkan nilai tiap matakuliah yang telah ditempuh mahasiswa.

Data nilai yang dapat diinput yaitu nim mahasiswa, nama mahasiswa, nama matakuliah, serta nilai yang diperoleh mahasiswa. Nilai yang telah diinput tersebut digunakan dalam melakukan perhitungan prediksi kelulusan mahasiswa. Hanya admin yang dapat mengakses menu ini.

C. Form Daftar Mahasiswa

Pada menu daftar mahasiswa merupakan menu yang berisi rekap nilai IPK yang diperoleh tiap mahasiswa selama masa studi dan total sks yang telah ditempuh. Menu ini terdapat tombol proses yang digunakan untuk mengetahui prediksi kelulusan mahasiswa. Pada kolom NB berisi daftar matakuliah yang harus ditingkatkan oleh mahasiswa agar dapat meningkatkan nilai IPK pada semester berikutnya dan dapat studinya tepat waktu.



Gambar 7 : Form Daftar Mahasiswa

D. Form Tambah Matakuliah



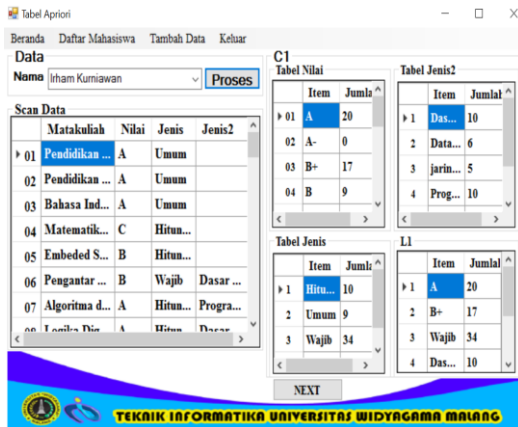
Gambar 8 : Form Tambah Matakuliah

Menu tambah matakuliah digunakan admin untuk menambahkan kode matakuliah, nama matakuliah, semester, sks, dan jenis matakuliah.

E. Form Apriori

Pada menu apriori dilakukan proses perhitungan C1 yaitu penggabungan jumlah item dari jenis matakuliah dan perolehan nilai dari matakuliah tersebut. L1 merupakan nilai support dari itemset 1 yang diperoleh dari hasil perhitungan C1.



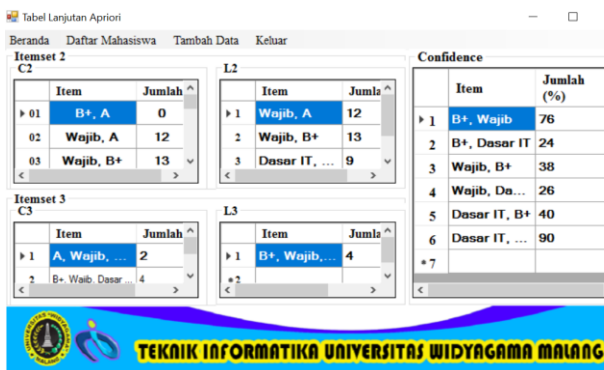


Gambar 9 : Form Apriori

- 1) Sistem ini dibuat untuk membantu mengetahui nilai matakuliah dan indeks prestasi kumulatif (IPK) yang digunakan untuk memprediksi kelulusan mahasiswa Teknik Informatika Universitas Widyagama.
- 2) Hasil perhitungan dengan algoritma apriori menunjukkan bahwa nilai *itemset* ke-3 dari data matakuliah mahasiswa merupakan nilai matakuliah dominan yang digunakan dalam meningkatkan indeks prestasi kumulatif pada semester berikutnya. Perolehan nilai tersebut dapat digunakan untuk memprediksi tahun lulus mahasiswa. Tabel hasil prediksi kelulusan adalah sebagai berikut :

Tabel 7 : Hasil Prediksi Kelulusan

| NO | NAMA                  | IPK  | TAHUN LULUS |
|----|-----------------------|------|-------------|
| 1  | Irham Kurniawan       | 3,55 | 2017        |
| 2  | Vibyola Putri W       | 3,72 | 2018        |
| 3  | Faruq Naufal Abdullah | 3,15 | 2019        |
| 4  | Muhammad Fahmi Imanda | 3,20 | 2020        |



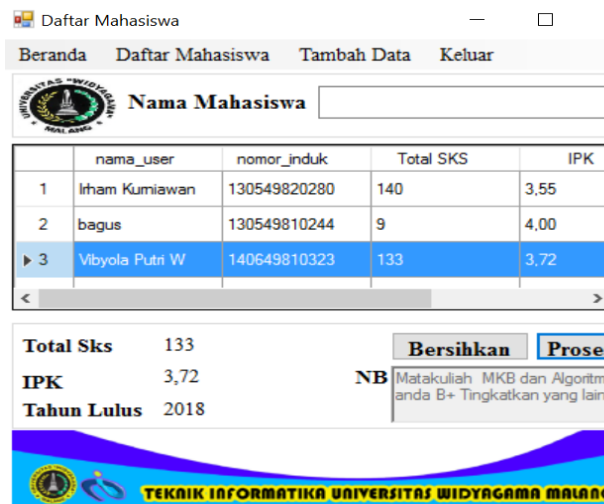
Gambar 10 : Form Lanjutan Apriori

**B. Saran**

- 1) Penelitian ini dapat digunakan sebagai acuan / referensi untuk menentukan seberapa cepat mahasiswa menyelesaikan studinya dan mengetahui matakuliah mana yang menjadi keunggulan dari mahasiswa tersebut.
- 2) Aplikasi ini hanya berjalan ketika mahasiswa tersebut ada di atas semester 5, karena matakuliah jenis 2 mulai semester 3.

Setelah itu dilanjutkan dengan perhitungan nilai *confidence*. Sehingga diperoleh *presentase* nilai dari *confidence* tiap matakuliah. Menu ini hanya dapat diakses oleh admin.

**F. Form Hasil Prediksi Kelulusan**



Gambar 11 : Form Hasil Prediksi Kelulusan

**V. REFERENSI**

- [1] Hartanto, D., & Hansun, S. (2014). IMPLEMENTASI DATA MINING DENGAN ALGORITMA C4.5 UNTUK MEMPREDIKSI TINGKAT KELULUSAN MAHASISWA. Jurnal ULTIMATICS , 1, 15-20. Dapat diakses di : <http://www.ejournals.umn.ac.id/index.php/TI/article/view/327.html> ].
- [2] Wirdasari, D. & A.Calam,.Penerapan Data Mining Untuk Mengolah Data Penempatan Buku Di Perpustakaan Smk Ti Pab 7 Lubuk Pakam Dengan Metode Association Rule. Jurnal SAINTIKOM, 10 (2), 150, 2011. Dapat diakses di : [http://www.academia.edu/7489155/Penerapan\\_Data\\_Mining\\_Untuk\\_Mengolah\\_Data\\_Penempatan\\_Buku\\_Di\\_Perpustakaan\\_Smk\\_Ti\\_Pab\\_7\\_Lubuk\\_Pakam\\_Dengan\\_Metode\\_Association\\_Rule](http://www.academia.edu/7489155/Penerapan_Data_Mining_Untuk_Mengolah_Data_Penempatan_Buku_Di_Perpustakaan_Smk_Ti_Pab_7_Lubuk_Pakam_Dengan_Metode_Association_Rule)].
- [3] Linoff , G. S., & Berry, M. J. *Data Mining Techniques for Marketing, Sales, Customer Relationship Management*. 2011. United States of America: Wiley Publishing, Inc. Dapat diakses di: <https://www.amazon.com/Data-Mining-Techniques-Relationship-Management/dp/0470650931>].

**V. KESIMPULAN**

**A. Simpulan**

Dari hasil penelitian yang telah dilakukan di Program Studi Teknik Informatika Universitas Widyagama Malang maka dapat disimpulkan bahwa:

- [4] David, Olson, & Yong, Shi. *Introduction to Business Data Mining*. 2011. International Edition: Mc Graw Hill. [\[https://journal.budiluhur.ac.id/index.php?journal=teleinformatika&page=index\]](https://journal.budiluhur.ac.id/index.php?journal=teleinformatika&page=index).
- [5] Han, J., & Kamber, M. *Data Mining: Concept and Techniques*. 2011. Waltham: Elsevier Inc.
- [6] Sensuse, G.G., Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan Pt. Gramedia. *Jurnal TELEMATIKA MKOM* , 4 (1), 118-132, 2012. Dapat diakses di :
- [7] Syaifullah, M. A. *Implementasi Data Mining Algoritma Apriori Pada Sistem Penjualan*. STMIK AMIKOM YOGYAKARTA. 2010. Yogyakarta: Sekolah Tinggi Manajemen Informatika dan Komputer Amikom. Dapat diakses di : [\[http://elearning.amikom.ac.id/index.php/karya/752/mardalina.%20se/Implementasi%20data%20mining%20algoritma%20apriori%20pada%20%20sistem%20penjualan\]](http://elearning.amikom.ac.id/index.php/karya/752/mardalina.%20se/Implementasi%20data%20mining%20algoritma%20apriori%20pada%20%20sistem%20penjualan).