

# Penerapan Data Mining Untuk Prediksi Daya Serap Lulusan Siswa Menggunakan Algoritma Native Bayes

Daka Waru<sup>1</sup>, Reny Wahyuning Astuti<sup>2</sup>, Novhirtamely Kahar<sup>3</sup>

<sup>123</sup>Fakultas Ilmu Komputer, Universitas Nurdin Hamzah Jambi, Indonesia

Email : [warudaka443@gmail.com](mailto:warudaka443@gmail.com), [r3ny4stuti@gmail.com](mailto:r3ny4stuti@gmail.com), [novmely@gmail.com](mailto:novmely@gmail.com)

## Article Information

### Article history

Received 25 April 2021

Revised 20 May 2021

Accepted 05 June 2021

Available 30 June 2021

### Keywords

Absorption  
Graduates  
Naive Bayes  
Rapidminer  
WEKA

### Corresponding Author:

Daka Waru,  
Fakultas Ilmu Komputer,  
Universitas Nurdin Hamzah Jambi,  
Indonesia,  
Email: [warudaka443@gmail.com](mailto:warudaka443@gmail.com)

## ABSTRACT

The importance of predicting the absorption of Vocational High School (SMK) graduates in the world of work, especially SMK Negeri 9 Muaro Jambi which is not yet known about the prediction of the world of work that accepts SMK graduates so that the purpose of this study is to analyze the prediction of the accuracy of the absorption of graduates of SMK Negeri 9 Muaro Jambi as material. a reference to see whether the graduates of SMK Negeri 9 Muaro Jambi have achieved the expected goals or not so that this analysis can be used as input for schools to improve the competence of SMK students. This implementation is assisted by using the Rapidminer and WEKA applications with 100 alumni work data input. The attributes used in this analysis process are Department, Waiting Time and Field of Work and Class of Work Field Accuracy. The process in this analysis is carried out with data that has been provided with the Naive Bayes Classification Method to predict the absorption of graduates. The results of this study the highest accuracy value in the Rapidminer application is at 100% and WEKA is at 100%.

**Keywords :** Absorption, Graduates, Naive Bayes, Rapidminer, WEKA

## ABSTRAK

Pentingnya dilakukan prediksi daya serap lulusan Sekolah Menengah Kejuruan (SMK) di dunia kerja terutama SMK Negeri 9 Muaro Jambi yang belum diketahui prediksi dunia kerja yang menerima lulusan SMK sehingga tujuan penelitian ini adalah untuk menganalisa prediksi ketepatan daya serap lulusan SMK Negeri 9 Muaro Jambi sebagai bahan acuan untuk melihat lulusan siswa SMK Negeri 9 Muaro Jambi apakah mencapai tujuan yang diharapkan atau tidak sehingga analisis ini dapat dijadikan sebagai masukan bagi sekolah untuk meningkatkan kompetensi siswa SMK. Implementasi ini dibantu dengan menggunakan aplikasi Rapidminer dan WEKA dengan input data pekerjaan alumni dengan jumlah 100 data. Atribut yang digunakan dalam proses analisis ini yaitu Jurusan, Lama Menunggu dan Bidang Pekerjaan serta Kelas Ketepatan Bidang Kerja. Proses dalam analisis ini dilakukan dengan data yang telah diberikan dengan Metode Klasifikasi Naive Bayes untuk memprediksi daya serap lulusan. Hasil dari penelitian ini nilai accuracy tertinggi pada aplikasi Rapidminer ada pada angka 100% dan WEKA ada pada angka 100%.

**Kata Kunci :** Absorption, Graduates, Naive Bayes, Rapidminer, WEKA

Copyright©2021 Daka Waru, Reny Wahyuning Astuti, Novhirtamely Kahar

This is an open access article under the [CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



## 1. Pendahuluan

Tantangan dunia kerja bagi Lulusan Sekolah Menengah Kejuruan perlu di perhatikan oleh pihak sekolah untuk meningkatkan kompetensi lulusan setiap siswa. Daya serap lulusan SMK oleh dunia kerja pada saat ini masih minim seiring dunia kerja terutama di perusahaan lebih memilih kompetensi dan keahlian para lulusan dalam bekerja sehingga dapat meningkatkan efisiensi dalam perusahaan tersebut.

Permasalahan yang muncul dilapangan pada Sekolah Menengah Kejuruan Negeri 9 Muaro Jambi bahwa masih terdapat beberap lulusan siswa yang belum mendapatkan pekerjaan setelah lulus sekolah. Sehingga sekolah perlu mengetahui data-data alumni yang belum mendaatkan pekerjaan tersebut serta berapa lama waktu tunggu para alumni mendapatkan pekerjaan di dunia kerja. Penelitian ini dianalisis dari data sekolah berdasarkan tahun terakhir.

Bagi sekolah sangat memerlukan data analisis untuk melihat serapan bagi lulusan siswa SMK yang telah diserap di dunia kerja, sehingga penelitian ini menggunakan metode *native bayes* sebagai algoritma untuk melakukan analisis dengan bantuan aplikasi rapid miner sebagai tools analisis datanya.

## 2. Kajian Terdahulu

Penelitian terdahulu bertujuan untuk mendapatkan bahan perbandingan dan acuan. Selain itu, untuk menghindari anggapan kesamaan dengan penelitian ini. Maka dalam kajian pustaka ini peneliti mencantumkan hasil-hasil penelitian terdahulu sebagai berikut:

### 1. Hasil Penelitian Kusuma

Pada Penelitian Kusuma L. N. (2019) bahwa penelitian ini menggunakan metode observasi dan studi kasus. Penelitian ini bertujuan untuk menganalisis bagaimana kemampuan lulusan SMK Buddhi Tangerang sehingga dapat bersaing di dunia kerja. Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa berdasarkan hasil evaluasi dan validasi, “diketahui bahwa algoritma Naïve Bayes memiliki akurasi dan performa baik dengan nilai accuracy sebesar 98% dan nilai AUC sebesar 0.980. Skor kesimpulan aplikasi berdasarkan SQA sebesar 84 yang dinilai cukup baik. Dengan keterangan dari 101 siswa diprediksi 51,48% siswa mampu bersaing dalam dunia kerja, sedangkan 48,52% siswa belum mampu bersaing dalam dunia kerja.”

### 2. Hasil Penelitian Shiddieq, D., & Fadillah, M

Pada Penelitian Shiddieq D. F. & Fadillah M. R. (2020) bahwa Penelitian ini merupakan penelitian yang menggunakan metode observasi dan studi kasus. Penelitian ini bertujuan untuk mengetahui informasi sudah sejauh mana keberhasilan pada setiap kompetensi keahlian dengan para lulusan yang sudah bekerja dengan posisi yang mereka dapatkan, apakah posisi yang mereka dapatkan sudah sesuai dengan kompetensi keahlian yang mereka miliki atau tidak sesuai.

### 3. Hasil Penelitian Midiarso, R., & Umilasari, R

Penelitian Midiarso, R., & Umilasari, R. (2019) bahwa Penelitian ini merupakan penelitian yang menggunakan metode observasi dan studi kasus. Penelitian ini bertujuan untuk mengetahui mengetahui hasil penerapan mengetahui hasil evaluasi kinerja metode Naive Bayes untuk klasifikasi status alumni pada SMK Bustanul Ulum Al-Ghazali. Berdasarkan penelitian yang dilakukan dapat disimpulkan bahwa “alumni yang bekerja sangat sesuai dengan bidang kompetensi keahlian sebesar 58 %, alumni yang bekerja cukup sesuai dengan bidang kompetensi keahlian sebesar 24 %, dan alumni yang bekerja tidak sesuai dengan kompetensi keahlian sebesar 18 %. Sehingga dengan adanya presentase tersebut pihak sekolah dapat mengetahui kesesuaian bidang pekerjaan alumni dengan jurusan yang diambil dengan akurat serta memudahkan pihak sekolah dalam melakukan evaluasi baik dalam kurikulum pembelajaran maupun kerja sama dengan perusahaan.”

### 3. Metodologi Penelitian

Penelitian ini dilakukan dengan menggunakan pendekatan atau metode studi kasus (*case research*) dan metode survei sengan pendekatan penelitian kuantitatif sehingga penelitian ini dapat dengan tepat mengukur ketepatan daya serap dunia kerja bagi lulusan SMK dalam kurun waktu tahun terakhir.

Analisis daya serap lapangan kerja bagi lulusan Sekolah Menengah Kejuruan (SMK) Negeri Muaro Jambi menggunakan algoritma Native Bayes. Pemilihan Algoritma Native Bayes memberikan keuntungan karena penggunaan Native Bayes tidak memerlukan data yang banyak seperti hal nya algoritma lain. Analisi ini yang nantinya akan didapatkan angka kuantitatif sebagai masukan bagi sekolah untuk meningkatkan kompetensi lulusan siswa SMK agar dapat bisa bersaing didunia kerja.

Tools yang digunakan untuk menganalisis data pada penelitian ini adalah menggunakan Aplikasi Rapid Miner dan WEKA sehingga diharapkan pengukuran dapat lebih cepat dan tepat untuk menghasilkan data output pengukuran.

#### 3.1 Kerangka Kerja Penelitian

Serangkaian proses pada tahapan penelitian ini akan dibuat berdasarkan tahapan yang terstruktur sehingga kerangka kerja pada penelitian ini dapat terukur dan terarah sesuai tahapannya. Adapun tahapan-tahapan kerangka kerja penelitian yang digunakan adalah sebagai berikut:



**Gambar 1 Kerangka Kerja Penelitian**

Berdasarkan kerangka kerja penelitian pada gambar 1 diatas, maka dapat diuraikan penjelasan masing-masing tahapan dalam penelitian adalah sebagai berikut :

#### 1. Identifikasi Masalah Pada SMK Negeri 9 Muaro Jambi

Pada tahap ini dilakukan identifikasi masalah yang terdapat pada SMK Negeri 9 Muaro Jambi. Bertujuan untuk menentukan permasalahan yang diangkat dalam penelitian berupa sebuah data informasi alumni atau lulusan yang bertujuan untuk mengembangkan data yang dirasakan kurang dioptimalkan lagi untuk dijadikan data yang dapat dimanfaatkan sebagai sumber data bagi sekolah untuk meningkatkan kompetensi lulusan siswa.

#### 2. Studi Literatur Tentang Naive Bayes dan Daya Serap Lulusan

Salah satu keuntungan menggunakan Native Bayes adalah tidak memerlukan data yang banyak pada sampel data sehingga ini dapat menghemat waktu serta dapat dengan cepat dalam melakukan analisis datanya. Pada tahap ini dilakukan pencarian landasan-landasan teori yang diperoleh dari berbagai buku, jurnal, dan internet untuk melengkapi konsep dan teori yang digunakan. Tujuannya agar teori yang dibahas memiliki landasan dan keilmuan yang ilmiah dari penelitian tersebut. Sehingga dapat mempermudah dalam memahami konsep dan teori yang digunakan dari penelitian yang dibahas.

#### 3. Pengumpulan Data Alumni

Pada tahap ini, penulis melakukan pengumpulan data dengan metode cara survei dan pengambilan data alumni atau lulusan dari SMK Negeri 9 Muaro Jambi Jurusan Multimedia yang didapatkan dari sumber data sekolah pada tahun terakhir

#### 4. Mengimplementasi Perangkat Lunak dengan RapidMiner dan Weka

Tahap implementasi dilakukan dengan aplikasi Microsoft Excel dan RapidMiner. Data mentah yang didapatkan dari sumber data primer yaitu data dari Sekolah akan di Ujikan secara manual diolah menggunakan excel kemudian akan di ujikan menggunakan Aplikasi RapidMiner dan Weka.

5. Menganalisis Kinerja Perangkat Lunak RapidMiner dan Weka

Tahap pengujian merupakan suatu teknik yang digunakan untuk menentukan bahwa implementasi yang dihasilkan telah mampu memecahkan masalah. Pada tahap ini, implementasi yang baru dikerjakan diuji kemampuan dan keefektifannya sehingga didapatkan kekurangan dan kelemahan implementasi yang kemudian dilakukan pengkajian ulang dan perbaikan terhadap implementasi menjadi lebih baik dan sempurna.

6. Menarik Kesimpulan

Langkah selanjutnya dalam penelitian ini adalah mengambil kesimpulan. Dalam mengambil kesimpulan perlu kehati – hatian agar tidak terjerumus kepada kesimpulan yang terlalu sempit atau terlalu luas, kesimpulan yang terlalu sempit bisa terjadi apabila ada ciri-ciri atau sifat atau fakta atau lainnya yang bersifat penting dalam ruang lingkup penelitian. Kesimpulan yang terlalu luas bisa terjadi apabila kesimpulan melebihi dari ruang lingkup penelitian. Suatu hal yang perlu diperhatikan adalah rumusan masalah atau pertanyaan penelitian, harus terjawab di dalam kesimpulan hasil penelitian

4. Hasil dan Pembahasan

4.1 Implementasi *Rapidminer*

Proses tahapan-tahapan implementasi ini akan menjelaskan bagaimana proses menggunakan aplikasi *Rapidminer* dengan metode *Naive Bayes* untuk mendapatkan prediksi daya serap lulusan

4.1.1 Tampilan Proses *Naive Bayes*

Pada tampilan hasil proses ini menunjukkan hasil olahan dengan menggunakan metode *Naive Bayes*. Dapat dilihat pada gambar-gambar berikut ini:

| accuracy: 100.00%  |            |                   |                  |                 |
|--------------------|------------|-------------------|------------------|-----------------|
|                    | true Tepat | true Kurang Tepat | true Tidak Tepat | class precision |
| pred. Tepat        | 16         | 0                 | 0                | 100.00%         |
| pred. Kurang Tepat | 0          | 5                 | 0                | 100.00%         |
| pred. Tidak Tepat  | 0          | 0                 | 9                | 100.00%         |
| class recall       | 100.00%    | 100.00%           | 100.00%          |                 |

Gambar 2 Tampilan *Table View* Ketepatan Kerja

Pada gambar di atas ditampilkan hasil perhitungan dari *Rapidminer* dalam bentuk tabel. Dengan keterangan *accuracy* 100%, *kappa* 1,00, *class recall* 100% pada Tepat, 100% pada Kurang Tepat dan 100% pada Tidak Tepat. Untuk *class precision*, 100% untuk Tepat, 100% untuk Kurang Tepat dan 100% untuk Tidak Tepat.



**PerformanceVector**

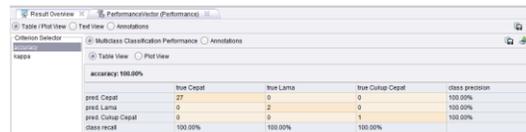
PerformanceVector:  
accuracy: 100.00%

ConfusionMatrix:  
True: Tepat Kurang Tepat Tidak Tepat  
Tepat: 16 0 0  
Kurang Tepat: 0 5 0  
Tidak Tepat: 0 0 9  
Kappa: 1.000

ConfusionMatrix:  
True: Tepat Kurang Tepat Tidak Tepat  
Tepat: 16 0 0  
Kurang Tepat: 0 5 0  
Tidak Tepat: 0 0 9

**Gambar 3** Tampilan *Text View* Ketepatan Kerja

Pada *confusion matrix* hasil perhitungan metode Naïve Bayes dengan RapidMiner, dari 30 data testing dapat diketahui bahwa dari 16 orang alumni yang diprediksi masuk kategori Tepat, 16 alumni masuk kategori Tepat. 5 orang alumni yang diprediksi masuk kategori Kurang Tepat, 5 alumni masuk kategori Kurang Tepat. 9 orang alumni yang diprediksi masuk kategori Tidak Tepat, 9 alumni masuk kategori Tidak Tepat. Dengan demikian, dari *confusion matrix* yang ada antara prediksi dan alumni yang masuk kategori akurasi 100%.



**Result Overview** PerformanceVector (Performance)

Criterion Selector  
Kappa

accuracy: 100.00%

|                  | true Cepat | true Lama | true Cukup Cepat | class precision |
|------------------|------------|-----------|------------------|-----------------|
| pred Cepat       | 27         | 0         | 0                | 100.00%         |
| pred Lama        | 0          | 2         | 0                | 100.00%         |
| pred Cukup Cepat | 0          | 1         | 100.00%          |                 |
| class recall     | 100.00%    | 100.00%   | 100.00%          |                 |

**Gambar 4** Tampilan *Table View* Waktu Tunggu

Pada gambar di atas ditampilkan hasil perhitungan dari Rapidminer dalam bentuk tabel. Dengan keterangan *accuracy* 100%, *kappa* 1,00, *class recall* 100% pada Cepat, 100% pada Cukup Cepat dan 100% pada Lama. Untuk *class precision*, 100% untuk Cepat, 100% untuk Cukup Cepat dan 100% untuk Lama.



**PerformanceVector**

PerformanceVector:  
accuracy: 100.00%

ConfusionMatrix:  
True: Cepat Lama Cukup Cepat  
Cepat: 27 0 0  
Lama: 0 2 0  
Cukup Cepat: 0 0 1  
Kappa: 1.000

ConfusionMatrix:  
True: Cepat Lama Cukup Cepat  
Cepat: 27 0 0  
Lama: 0 2 0  
Cukup Cepat: 0 0 1

**Gambar 5** Tampilan *Text View* Waktu Tunggu

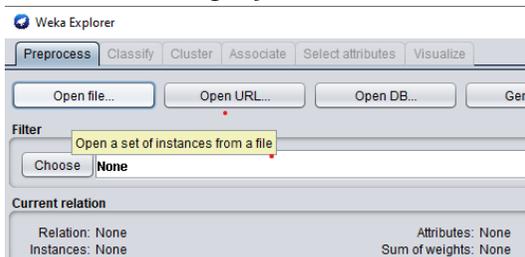
Pada *confusion matrix* hasil perhitungan metode Naïve Bayes dengan RapidMiner, dari 30 data testing dapat diketahui bahwa dari 27 orang alumni yang diprediksi masuk kategori Cepat, 27 alumni masuk kategori Cepat. 1 orang alumni yang diprediksi masuk kategori Cukup Cepat, 1 alumni masuk kategori Cukup Cepat. 2 orang alumni yang diprediksi masuk kategori Lama, 1 alumni masuk kategori Lama. Dengan demikian, dari *confusion matrix* yang ada antara prediksi dan alumni yang masuk kategori akurasi 100%.

## 4.2 Implementasi WEKA

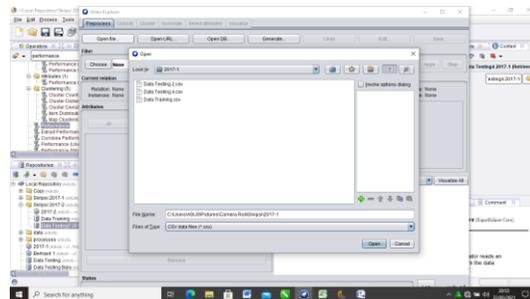
Proses tahapan-tahapan implementasi ini akan menjelaskan bagaimana proses menggunakan aplikasi WEKA dengan metode *Naive Bayes* untuk mendapatkan prediksi daya serap lulusan yang telah dilakukan peneliti. Adapun proses tahapannya adalah sebagai berikut:

### 4.2.1 Proses Perhitungan

Pada proses ini dilakukan proses perhitungan, mulai dari memasukkan data, memilih metode, hingga dapat diketahui hasilnya. Pada awalnya file dengan tipe .csv akan dibuka dengan menekan tombol *open file*.

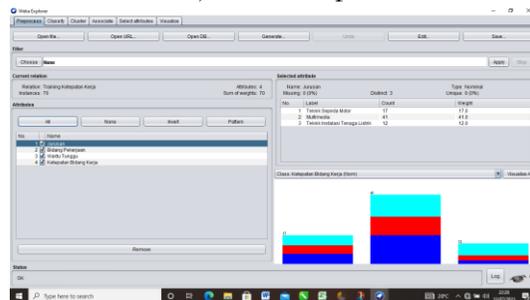


Gambar 6 Tombol perintah open file



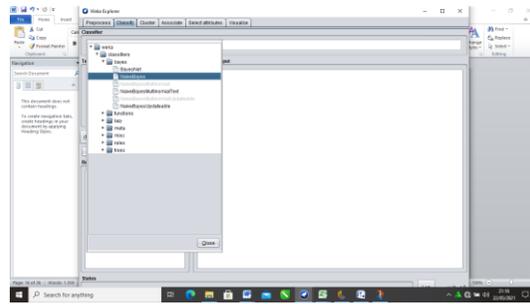
Gambar 7 Pemilihan file yang akan di-open

Pada gambar dibawah ini menunjukkan tampilan WEKA setelah membuka file.



Gambar 8 Tampilan setelah file di-open

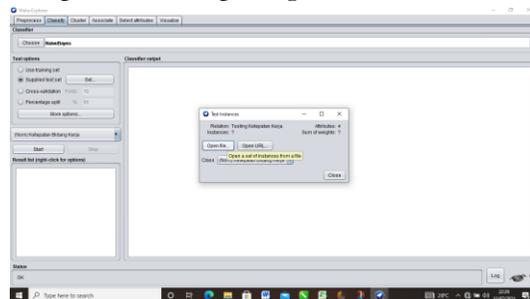
Setelah membuka data, kemudian harus memilih metode yang akan digunakan. Untuk *naïve bayes* metode dapat di pilih dalam sub-menu *classify*.



Gambar 9 pemilihan metode *Naive Bayes*

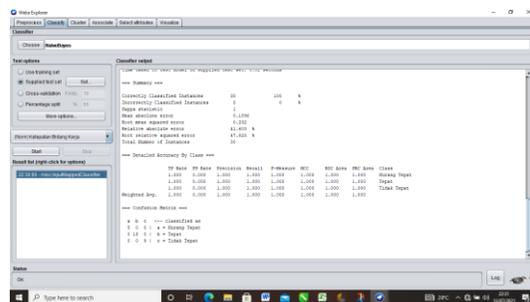
Setelah mengikuti langkah dari gambar diatas kemudian memilih data test yang akan digunakan.

Pada proses perhitungan dengan *Naive Bayes* dapat dilakukan dengan atau tanpa data test, tapi pada penelitian kali ini, peneliti menggunakan perhitungan dengan menggunakan data test, seperti terlihat pada gambar berikut ini :



Gambar 10 open file data test

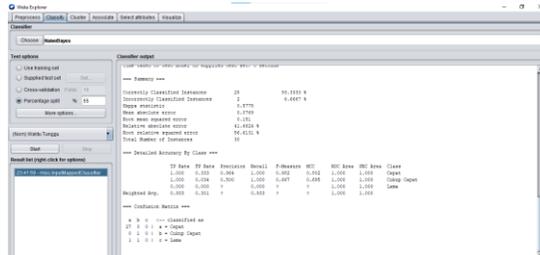
Setelah *data testing di-input*, tekan tombol start dan WEKA akan otomatis melakukan perhitungan.



Gambar 11 Tampilan hasil perhitungan Ketepatan Kerja

Pada gambar di atas ditampilkan hasil perhitungan dari WEKA dalam bentuk teks. Dengan keterangan *accuracy 100%*, *kappa statistic 1*, *class recall 1,00* pada Tepat, 1,00 pada Kurang Tepat dan 1,00 pada Tidak Tepat. Untuk *class precision*, 1,00 untuk Tepat, 1,00 untuk Kurang Tepat dan 1,00 untuk Tidak Tepat.

Dari *confusion matrix* hasil perhitungan metode Naïve Bayes dengan WEKA dan *test option Supplied Test Set*, dari 30 data testing dapat diketahui bahwa dari 5 orang alumni yang diprediksi masuk kategori Kurang Tepat, 5 alumni masuk kategori Kurang Tepat. 16 orang alumni yang diprediksi masuk kategori Tepat, 16 alumni masuk kategori Tepat. 9 orang alumni yang diprediksi masuk kategori Tidak Tepat, 9 alumni masuk kategori Tidak Tepat. Dengan demikian, dari *confusion matrix* yang ada antara prediksi dan alumni yang masuk kategori akurasi 100%.



Gambar 12 Tampilan hasil perhitungan Waktu Tunggu

Pada gambar di atas ditampilkan hasil perhitungan dari WEKA dalam bentuk teks. Dengan keterangan *accuracy* 93,33%, *kappa statistic* 0,5775, *class recall* 1,00 pada Cepat, 1,00 pada Cukup Cepat dan 0 pada Lama. Untuk *class precision*, 0,964 untuk Cepat, 0,500 untuk Cukup Cepat dan 0 untuk Lama.

Dari *confusion matrix* hasil perhitungan metode Naïve Bayes dengan WEKA dan *test option Supplied Test Set*, dari 30 data testing dapat diketahui bahwa dari 27 orang alumni yang diprediksi masuk kategori Cepat, 27 alumni masuk kategori Kurang Tepat. 1 orang alumni yang diprediksi masuk kategori Cukup Cepat, 1 alumni masuk kategori Cukup Cepat. 2 orang alumni yang diprediksi masuk kategori Lama, 1 alumni masuk kategori Cepat dan 1 alumni masuk kategori Cukup Cepat. Dengan demikian, dari *confusion matrix* yang ada antara prediksi dan alumni yang masuk kategori akurasi 93,33%.

### 4.3 Implementasi Microsoft Excel

Untuk mengetahui ketepatan hasil dari perhitungan prediksi daya serap lulusan, peneliti melakukan perbandingan dengan melakukan perhitungan menggunakan Microsoft Excel, dengan hasil sebagai berikut :

#### 4.3.1 Implementasi Ketepatan Kerja

##### 4.3.1.1 Persiapan Data Training

Data training diperoleh dari google form yang diisi oleh alumni, dari data diperoleh oleh peneliti, peneliti hanya menggunakan atribut yang diperlukan untuk perhitungan prediksi daya serap lulusan. Data training yang digunakan dalam penelitian kali ini sebanyak 70 record, yaitu sebagai berikut :

Tabel 1 Data Training Ketepatan Kerja

| No. | Jurusan                         | Bidang Pekerjaan        | Waktu Tunggu   | Ketepatan Bidang Kerja |
|-----|---------------------------------|-------------------------|----------------|------------------------|
| 1   | Teknik Sepeda Motor             | Melanjutkan Studi       | Tidak Menunggu | Kurang Tepat           |
| 2   | Multimedia                      | Wirausaha               | 1-6 Bulan      | Tepat                  |
| 3   | Multimedia                      | Polri                   | 1-6 Bulan      | Tidak Tepat            |
| 4   | Multimedia                      | Seni Kreatif dan Desain | Tidak Menunggu | Tepat                  |
| 5   | Multimedia                      | Mekanik                 | 1-6 Bulan      | Tidak Tepat            |
| 6   | Teknik Instalasi Tenaga Listrik | Melanjutkan Studi       | Tidak Menunggu | Kurang Tepat           |
| 7   | Multimedia                      | Melanjutkan Studi       | Tidak Menunggu | Kurang Tepat           |
| 8   | Teknik Sepeda Motor             | Melanjutkan Studi       | 1-6 Bulan      | Kurang Tepat           |
| 9   | Multimedia                      | Polri                   | 7-12 Bulan     | Tidak Tepat            |
| 10  | Multimedia                      | Mekanik                 | 1-6 Bulan      | Tidak Tepat            |
| 11  | Multimedia                      | Seni Kreatif dan Desain | Tidak Menunggu | Tepat                  |
| 12  | Multimedia                      | Wirausaha               | Tidak Menunggu | Tepat                  |
| ... | ...                             | ...                     | ...            | ...                    |
| 69  | Multimedia                      | Sales dan Marketing     | 7-12 Bulan     | Tidak Tepat            |
| 70  | Teknik Sepeda Motor             | Keamanan                | 1-6 Bulan      | Tidak Tepat            |

#### 4.3.1.2 Perhitungan Probabilitas

Perhitungan probabilitas dilakukan kepada semua kolom yang ada pada tabel, sehingga menghasilkan gambar-gambar berikut :

Tabel 2 Probabilitas Kolom Jurusan

| P(Jurusan=    ...               | Tepat  | Kurang Tepat | Tidak Tepat |
|---------------------------------|--------|--------------|-------------|
| Multimedia                      | 57,89% | 68,00%       | 50,00%      |
| Teknik Instalasi Tenaga Listrik | 15,79% | 8,00%        | 26,92%      |
| Teknik Sepeda Motor             | 26,32% | 24,00%       | 23,08%      |

Tabel 3 Probabilitas Kolom Bidang Pekerjaan

| P(Bidang Pekerjaan=    ... | Tepat  | Kurang Tepat | Tidak Tepat |
|----------------------------|--------|--------------|-------------|
| Melanjutkan Studi          | 0,00%  | 100,00%      | 0,00%       |
| Seni Kreatif dan Desain    | 21,05% | 0,00%        | 3,85%       |
| Teknisi Sepeda Motor       | 15,79% | 0,00%        | 0,00%       |
| Teknisi Listrik            | 10,53% | 0,00%        | 0,00%       |
| Wirausaha                  | 52,63% | 0,00%        | 0,00%       |
| Mekanik                    | 0,00%  | 0,00%        | 15,38%      |
| Sales dan Marketing        | 0,00%  | 0,00%        | 34,62%      |
| Buruh                      | 0,00%  | 0,00%        | 26,92%      |
| Polri                      | 0,00%  | 0,00%        | 11,54%      |
| Keamanan                   | 0,00%  | 0,00%        | 7,69%       |

Tabel 4 Probabilitas Kolom Lama Menunggu

| P(Lama Menunggu=    ... | Tepat | Kurang Tepat | Tidak Tepat |
|-------------------------|-------|--------------|-------------|
| Tidak Menunggu          | 32%   | 36%          | 12%         |
| 1-6 Bulan               | 58%   | 52%          | 58%         |
| 7-12 Bulan              | 11%   | 8%           | 27%         |
| 1-2 Tahun               | 0%    | 4%           | 4%          |

Tabel 5 Probabilitas Label/Kelas

| Tepat  | Kurang Tepat | Tidak Tepat |
|--------|--------------|-------------|
| 0,2714 | 0,3571       | 0,3714      |

#### 4.3.1.3 Hasil Perhitungan *Microsoft Excel*

Pada penelitian kali ini, peneliti menggunakan 30 data testing yang digunakan untuk memprediksi daya serap lulusan siswa SMK negeri 9 Muaro Jambi. Adapun contoh data testing dan hasil perhitungan dengan Naïve Bayes menggunakan Microsoft Excel dapat dilihat pada gambar dibawah ini.

Tabel 6 Contoh Hasil Perhitungan Data Testing

| No. | Jurusan                         | Bidang Pekerjaan        | Waktu Tunggu   | Ketepatan Bidan Kerja | CLASS PREDICTION | Tepat  | Kurang Tepat | Tidak Tepat |
|-----|---------------------------------|-------------------------|----------------|-----------------------|------------------|--------|--------------|-------------|
| 1   | Teknik Sepeda Motor             | Teknik Sepeda Motor     | 1-6 Bulan      | Tepat                 | Tepat            | 0,0065 | 0,0000       | 0,0000      |
| 2   | Multimedia                      | Wirasaha                | 1-6 Bulan      | Tepat                 | Tepat            | 0,0479 | 0,0000       | 0,0000      |
| 3   | Multimedia                      | Melanjutkan Studi       | 1-6 Bulan      | Kurang Tepat          | Kurang Tepat     | 0,0000 | 0,1263       | 0,0000      |
| 4   | Multimedia                      | Seni Kreatif dan Desain | 1-6 Bulan      | Tepat                 | Tepat            | 0,0192 | 0,0000       | 0,0041      |
| 5   | Multimedia                      | Melanjutkan Studi       | 1-2 Tahun      | Kurang Tepat          | Kurang Tepat     | 0,0000 | 0,0092       | 0,0000      |
| 6   | Multimedia                      | Melanjutkan Studi       | 1-6 Bulan      | Kurang Tepat          | Kurang Tepat     | 0,0000 | 0,1263       | 0,0000      |
| 7   | Multimedia                      | Seni Kreatif dan Desain | 1-6 Bulan      | Tepat                 | Tepat            | 0,0192 | 0,0000       | 0,0041      |
| 29  | Multimedia                      | Sales dan Marketing     | Tidak Menunggu | Tidak Tepat           | Tidak Tepat      | 0,0000 | 0,0000       | 0,0074      |
| 30  | Teknik Instalasi Tenaga Listrik | Teknisi Listrik         | Tidak Menunggu | Tepat                 | Tepat            | 0,0014 | 0,0000       | 0,0000      |

Tabel 7 Hasil Prediksi dan Tingkat Akurasi

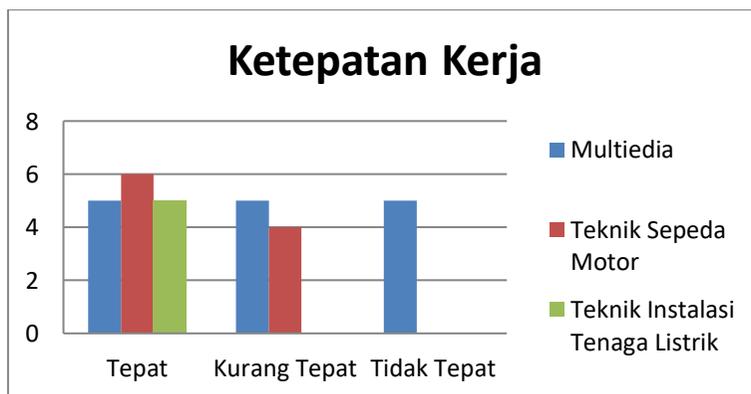
| 100%         | Predicted | CLASS |              |             |
|--------------|-----------|-------|--------------|-------------|
|              |           | Tepat | Kurang Tepat | Tidak Tepat |
| Tepat        |           | 16    | 0            | 0           |
| Kurang Tepat |           | 0     | 5            | 0           |
| Tidak Tepat  |           | 0     | 0            | 9           |

Pada *confusion table* hasil perhitungan metode Naïve Bayes dengan RapidMiner, dari 30 data testing dapat diketahui bahwa dari 16 orang alumni yang diprediksi masuk kategori Tepat, 16 alumni masuk kategori Tepat. 5 orang alumni yang diprediksi masuk kategori Kurang Tepat, 5 alumni masuk kategori Kurang Tepat. 9 orang alumni yang diprediksi masuk kategori Tidak Tepat, 9 alumni masuk kategori Tidak Tepat. Dengan demikian, dari *confusion table* yang ada antara prediksi dan alumni yang masuk kategori akurasi 100%.

Tabel 8 Tabel Perbandingan Hasil Aplikasi

| PERBANDINGAN    | RapidMiner | WEKA | Microsoft Excel |
|-----------------|------------|------|-----------------|
| <i>Accuracy</i> | 100%       | 100% | 100%            |
| <i>Kappa</i>    | 1.00       | 1    | -               |

Dari tabel perbandingan perhitungan diatas, dapat disimpulkan bahwa hasil perhitungan dari aplikasi WEKA dan RapidMiner dibandingkan dengan perhitungan Microsoft Excel memiliki hasil *accuracy* yang sama, yaitu 100% dan *kappa* 1,00.



Gambar 13 Grafik perbandingan Jurusan dan Ketepatan Kerja

Dari 30 data testing, 15 alumni merupakan jurusan Multimedia dan 5 alumni masuk kategori Tepat, 5 alumni kategori Kurang Tepat dan 5 alumni kategori Tidak Tepat. 10 alumni merupakan jurusan Teknik Sepeda Motor dan 6 alumni masuk kategori Tepat dan 4 alumni masuk kategori Kurang Tepat. 5 Alumni merupakan jurusan Teknik Instalasi Tenaga Listrik dan 5 alumni masuk kategori Tepat.

### 4.3.2 Implementasi Waktu Tunggu

#### 4.3.2.1 Persiapan Data Training

Data training diperoleh dari google form yang diisi oleh alumni, dari data diperoleh oleh peneliti, peneliti hanya menggunakan atribut yang diperlukan untuk perhitungan prediksi daya serap lulusan. Data training yang digunakan dalam penelitian kali ini sebanyak 70 record, yaitu sebagai berikut :

**Tabel 9 Data Training Waktu Tunggu**

| No. | Jurusan                         | Bidang Pekerjaan        | Lama Menunggu  | waktu Tunggu |
|-----|---------------------------------|-------------------------|----------------|--------------|
| 1   | Teknik Sepeda Motor             | Melanjutkan Studi       | Tidak Menunggu | Cepat        |
| 2   | Multimedia                      | Wirausaha               | 1-6 Bulan      | Cepat        |
| 3   | Multimedia                      | Polri                   | 1-6 Bulan      | Cepat        |
| 4   | Multimedia                      | Seni Kreatif dan Desain | Tidak Menunggu | Cepat        |
| 5   | Multimedia                      | Mekanik                 | 1-6 Bulan      | Cepat        |
| 6   | Teknik Instalasi Tenaga Listrik | Melanjutkan Studi       | Tidak Menunggu | Cepat        |
| 7   | Multimedia                      | Melanjutkan Studi       | Tidak Menunggu | Cepat        |
| 8   | Teknik Sepeda Motor             | Melanjutkan Studi       | 1-6 Bulan      | Cepat        |
| 9   | Multimedia                      | Polri                   | 7-12 Bulan     | Cukup Cepat  |
| 10  | Multimedia                      | Mekanik                 | 1-6 Bulan      | Cepat        |
| 11  | Multimedia                      | Seni Kreatif dan Desain | Tidak Menunggu | Cepat        |
| 12  | Multimedia                      | Wirausaha               | Tidak Menunggu | Cepat        |
| 69  | Multimedia                      | Sales dan Marketing     | 7-12 Bulan     | Cukup Cepat  |
| 70  | Teknik Sepeda Motor             | Keamanan                | 1-6 Bulan      | Cepat        |

#### 4.3.2.2 Perhitungan Probabilitas

Perhitungan probabilitas dilakukan kepada semua kolom yang ada pada tabel, sehingga menghasilkan gambar-gambar berikut :

**Tabel 10 Probabilitas Kolom Jurusan**

| P(Jurusan=   ...                | Cepat         | Cukup Cepat   | Lama          |
|---------------------------------|---------------|---------------|---------------|
| Multimedia                      | <b>59,65%</b> | <b>54,55%</b> | <b>50,00%</b> |
| Teknik Instalasi Tenaga Listrik | <b>15,79%</b> | <b>18,18%</b> | <b>50,00%</b> |
| Teknik Sepeda Motor             | <b>24,56%</b> | <b>27,27%</b> | <b>0,00%</b>  |

**Tabel 10 Probabilitas Kolom Bidang Pekerjaan**

| P(Bidang Pekerjaan=   ... | Cepat         | Cukup Cepat   | Lama          |
|---------------------------|---------------|---------------|---------------|
| Melanjutkan Studi         | <b>38,60%</b> | <b>18,18%</b> | <b>50,00%</b> |
| Seni Kreatif dan Desain   | <b>7,02%</b>  | <b>0,00%</b>  | <b>50,00%</b> |
| Teknisi Sepeda Motor      | <b>5,26%</b>  | <b>0,00%</b>  | <b>0,00%</b>  |
| Teknisi Listrik           | <b>1,75%</b>  | <b>9,09%</b>  | <b>0,00%</b>  |
| Wirausaha                 | <b>15,79%</b> | <b>9,09%</b>  | <b>0,00%</b>  |
| Mekanik                   | <b>7,02%</b>  | <b>0,00%</b>  | <b>0,00%</b>  |
| Sales dan Marketing       | <b>10,53%</b> | <b>27,27%</b> | <b>0,00%</b>  |
| Buruh                     | <b>7,02%</b>  | <b>27,27%</b> | <b>0,00%</b>  |
| Polri                     | <b>3,51%</b>  | <b>9,09%</b>  | <b>0,00%</b>  |
| Keamanan                  | <b>3,51%</b>  | <b>0,00%</b>  | <b>0,00%</b>  |

Tabel 11 Probabilitas Kolom Lama Menunggu

| P(Lama Menunggu=   ...) | Cepat | Cukup Cepat | Lama |
|-------------------------|-------|-------------|------|
| Tidak Menunggu          | 32%   | 0%          | 0%   |
| 1-6 Bulan               | 68%   | 0%          | 0%   |
| 7-12 Bulan              | 0%    | 100%        | 0%   |
| 1-2 Tahun               | 0%    | 0%          | 100% |

Tabel 12 Probabilitas Label/Kelas

| Cepat  | Cukup Cepat | Lama   |
|--------|-------------|--------|
| 0,8143 | 0,1571      | 0,0286 |

### 4.3.2.3 Hasil Perhitungan Microsoft Excel

Pada penelitian kali ini, peneliti menggunakan 30 data testing yang digunakan untuk memprediksi daya serap lulusan siswa SMK negeri 9 Muaro Jambi. Adapun contoh data testing dan hasil perhitungan dengan Naïve Bayes menggunakan Microsoft Excel dapat dilihat pada gambar dibawah ini.

Tabel 13 Contoh Hasil Perhitungan Data Testing

| No | Kategori                        | Label/Kelas           | Lama Menunggu | Waktu Menunggu | CLASS       | Probabilitas | Cepat  | Cukup Cepat | Lama   |
|----|---------------------------------|-----------------------|---------------|----------------|-------------|--------------|--------|-------------|--------|
| 1  | Siapa-Siapa Motor               | Siapa-Siapa Motor     | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 2  | Multimedia                      | Multimedia            | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 3  | Multimedia                      | Multimedia            | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 4  | Multimedia                      | Servisruff dan Desain | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 5  | Multimedia                      | Multimedia            | 1-2 Tahun     | Lama           | Lama        | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 6  | Multimedia                      | Multimedia            | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 7  | Multimedia                      | Servisruff dan Desain | 1-6 Bulan     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 28 | Multimedia                      | Sales dan Marketing   | 7-12 Bulan    | Cukup Cepat    | Cukup Cepat | 0,0000       | 0,0000 | 0,0000      | 0,0000 |
| 29 | Teknik Instalasi Tenaga Listrik | Teknik Listrik        | 1-2 Tahun     | Cepat          | Cepat       | 0,0000       | 0,0000 | 0,0000      | 0,0000 |

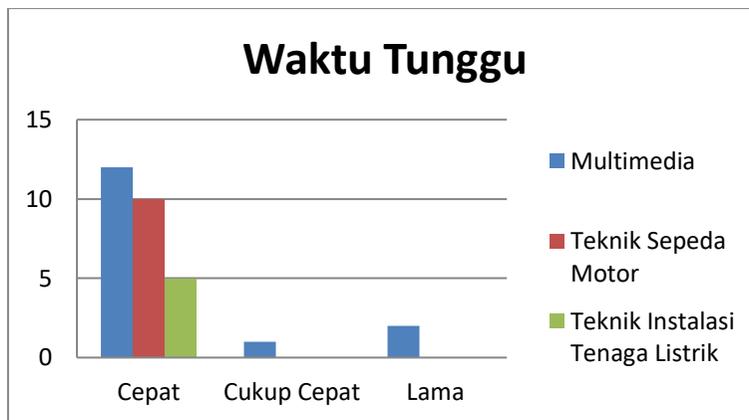
Tabel 14 Hasil Prediksi dan Tingkat Akurasi

| 100% | Predicted   | CLASS |             |      |
|------|-------------|-------|-------------|------|
|      |             | Cepat | Cukup Cepat | Lama |
|      | Cepat       | 27    | 0           | 0    |
|      | Cukup Cepat | 0     | 1           | 0    |
|      | Lama        | 0     | 0           | 2    |

Pada *confusion table* hasil perhitungan metode Naïve Bayes dengan RapidMiner, dari 30 data testing dapat diketahui bahwa dari 27 orang alumni yang diprediksi masuk kategori Cepat, 27 alumni masuk kategori Cepat. 1 orang alumni yang diprediksi masuk kategori Cukup Cepat, 1 alumni masuk kategori Cukup Cepat. 2 orang alumni yang diprediksi masuk kategori Lama, 1 alumni masuk kategori Lama. Dengan demikian, dari *confusion table* yang ada antara prediksi dan alumni yang masuk kategori akurasi 100%.

Tabel 15 Tabel Perbandingan Hasil Aplikasi

| PERBANDINGAN    | RapidMiner | WEKA   | Microsoft Excel |
|-----------------|------------|--------|-----------------|
| <i>Accuracy</i> | 100%       | 93,33% | 100%            |
| <i>Kappa</i>    | 1.00       | 0,5775 | -               |



Gambar 14 Grafik perbandingan Jurusan dan Waktu Tunggu

Dari 30 data testing, 15 alumni merupakan jurusan Multimedia dan 12 alumni masuk kategori Cepat, 1 alumni kategori cukup cepat dan 2 alumni kategori lama. 10 alumni merupakan jurusan Teknik Sepeda Motor dan 10 alumni masuk kategori Cepat. 5 Alumni merupakan jurusan Teknik Instalasi Tenaga Listrik dan 5 alumni masuk kategori Cepat.

Dari tabel perbandingan perhitungan diatas, dapat disimpulkan bahwa hasil perhitungan dari aplikasi WEKA dan RapidMiner dibandingkan dengan perhitungan Microsoft Excel memiliki hasil *accuracy* yang berbeda, yaitu 100% untuk RapidMiner dan Microsoft Excel serta 93,33% untuk Weka.

## 5. Kesimpulan

Setelah peneliti melakukan analisis *data mining* daya serap lulusan yang bertujuan untuk memprediksi daya serap lulusan SMK Negeri 9 Muaro Jambi, peneliti menarik kesimpulan bahwa

Pengujian algoritma *Naive Bayes* dengan *Rapidminer* dan *WEKA* menunjukkan bahwa perhitungan manual dibandingkan dengan perhitungan aplikasi memiliki hasil yang sama yaitu 100%. Sementara pengujian menunjukkan bahwa persentase dari prediksi bisa saja sama, namun hasil *confusion table* atau *confusion matrix* yang berbeda.

Nilai *accuracy* pada aplikasi *Rapidminer* untuk perhitungan Ketepatan Kerja ada pada angka 100% dan untuk perhitungan Waktu Tunggu ada pada angka 100% dimana Nilai *accuracy* pada aplikasi *WEKA* untuk perhitungan Ketepatan Kerja ada pada angka 100% dan untuk perhitungan Waktu Tunggu ada pada angka 93,33%.

Untuk Perhitungan Waktu Tunggu, dari 30 *data record*, Jurusan Multimedia menghasilkan 15 alumni dan 12 alumni masuk kategori Cepat, 1 alumni kategori Cukup Cepat dan 2 alumni kategori Lama. Jurusan Teknik Sepeda Motor menghasilkan 10

alumni dan 10 alumni masuk kategori Cepat. Jurusan Teknik Instalasi Tenaga Listrik menghasilkan 5 alumni dan 5 alumni masuk kategori Cepat.

Dari 30 *data record*, Jurusan Multimedia menghasilkan 15 alumni dan 5 alumni masuk kategori Tepat, 5 alumni kategori Kurang Tepat dan 5 alumni kategori Tidak Tepat. Jurusan Teknik Sepeda Motor menghasilkan 10 alumni dan 6 alumni masuk kategori Tepat, 4 alumni masuk kategori Kurang Tepat. Jurusan Teknik Instalasi Tenaga Listrik menghasilkan 5 alumni dan 5 alumni masuk kategori Tepat.

## 6. Ucapan Terima Kasih

Terima kasih kepada Universitas Nurdin Hamzah Jambi, Fakultas Ilmu Komputer serta SMK Negeri 9 Muaro Jambi yang telah mengizinkan melakukan penelitian di SMK Negeri 9 Muaro Jambi.

## 7. Pernyataan Penulis

Saya Daka Waru menyatakan bahwa penelitian yang berjudul “Data Mining Prediksi Profil Daya Serap Lulusan Siswa Smk Negeri 9 Muaro Jambi” adalah benar-benar hasil penelitian dan penulisan sendiri bukan hasil plagiat dan atau saduran dan atau jiplakan dari hasil penelitian orang lain, dan apabila dikemudian hari pernyataan ini tidak benar, maka saya bersedia menerima sanksi dari peraturan yang berlaku.

## Bibliografi

- Alan, Dennis Et Al. 2013. *System Analysis And Design With Uml 4<sup>th</sup> Edition*. John Wiley And Sons.
- Anhar, 2020, “Ilmu Komputer”, Dilihat 18 September 2020, <https://ilmukomputer.org/wp-content/uploads/2009/06/Anharku-Flowchart.Pdf>
- Astuti, Reny., Puspitorini, S., & Akbar, Fajri. 2020. Implementasi Teorema Naïve Bayes Pada Analisa Dan Prediksi Bidang Pekerjaan Alumni Prodi Teknik Informatika Stmik Nurdin Hamzah Jambi. *Fortech (Journal Of Information Technology)* Vol. 4 No. 1.
- Eko Prasetyo, 2012. *Data Mining Konsep Dan Aplikasi Menggunakan Matlab*. Andi, Yogyakarta.
- Kusuma, L. 2019. Prediksi Kemampuan Lulusan Smk Untuk Dapat Bersaing Di Dunia Kerja Dengan Menggunakan Naïve Bayes: Studi Kasus Smk Buddhi Tangerang. *Algor*,1(1) 56-63.

- Manalu, E., Sianturi, F. A., & Manalu, M. R. 2017. Penerapan Algoritma Naïve Bayes Untuk Memprediksi Jumlah Produksi Barang Berdasarkan Data Persediaan Dan Jumlah Pemesanan Pada Cv. Papadan Mama Pastries. *Jurnal Mantik Penusa*, 1(2).
- Midiarso, R., & Umilasari, R. 2019. Pemanfaatan Algoritma Naive Bayes Untuk Klasifikasi Status Alumni Smk Bustanul Ulum Al-Ghazali Wuluhan Jember. *Undergraduate Thesis*. Universitas Muhammadiyah Jember.
- Nursubiyantoro, Eko & Puryani, Puryani. 2016. Perancangan Sistem Penelusuran Alumni (Tracer Study) Berbasis Web. *Opsi*. 9. 85. 10.31315/Opsi.V9i2.2228.
- Shiddieq, D., & Fadillah, M. 2020. Penerapan Metode K-Means Untuk Klasifikasi Bidang Pekerjaan Alumni. *Jurnal Komputer Bisnis*, 12(2), 58-62.
- Suliyanto, 2012. “Analisis Statistik Pendekatan Praktis Dengan Microsoft Excel”, Andi, Yogyakarta.
- Suyanto. 2017. *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Bandung, Informatika.
- Tri Vulandari, R., 2017. *Data Mining Teori Dan Aplikasi Rapidminer*, Penerbit Gava Media, Yogyakarta.
- Vijayarani, D., & Dhayanand, M. 2015. *Liver Disease Prediction Using Svm And Naïve Bayes Algorithms*. India: International Journal Ofscience, Engineering And Technology Research.