

# Penerapan Data Mining dengan Metode Klasifikasi Naïve Bayes untuk Menilai Kinerja Asesor pada BAN PAUD dan PNF Provinsi Sumatera Utara

Nurhayatia<sup>1</sup>, Edy Victor Haryantob<sup>2</sup>, Linda Wahyuni<sup>3</sup>, Noprita Elisabeth S<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Sistem Informasi UPU, K.L. Yos Sudarso KM. 6,5 Tj. Mulia No. 3A, Medan

e-mail: [havatipearce66@gmail.com](mailto:havatipearce66@gmail.com)\*<sup>1</sup>, [edyvictor@gmail.com](mailto:edyvictor@gmail.com)<sup>2</sup>, [lindawahyuni391@gmail.com](mailto:lindawahyuni391@gmail.com)<sup>3</sup>, [novrvelisa@gmail.com](mailto:novrvelisa@gmail.com)<sup>4</sup>

## INFORMASI ARTIKEL

### Sejarah Artikel:

Diterima Redaksi : 2 Maret 2024

Revisi Akhir : 15 Mei 2024

Diterbitkan Online : 30 Mei 2024

### Kata Kunci:

Penilaian Kinerja, Asesor, Naïve Bayes, Weka.

### Korespondensi :

Telepon / Hp : +62 (0265) 272727

E-mail : [email@afiliasi.ac.id](mailto:email@afiliasi.ac.id)

## A B S T R A K

Akreditasi satuan Pendidikan Anak Usia Dini (PAUD) dan Pendidikan Nonformal (PNF) merupakan suatu rangkaian kegiatan yang dilakukan secara bertahap, terencana dan terukur untuk menilai kelayakan dan mutu dari satuan PAUD dan PNF sebagaimana telah diatur dalam Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. Dalam pelaksanaan kegiatannya, akreditasi dilakukan berdasarkan pada ketentuan yang diatur oleh Badan Akreditasi Nasional Pendidikan Anak Usia Dini dan Pendidikan Nonformal (BAN PAUD dan PNF). Adapun proses dalam penilaian tersebut dilakukan oleh Asesor. Pentingnya memiliki asesor yang berkualitas dan kompeten untuk melakukan penilaian dan akreditasi terhadap lembaga PAUD dan PNF di Provinsi Sumatera Utara. Namun, tidak semua asesor memiliki kompetensi dan kualitas yang sama dalam melakukan penilaian dan akreditasi. Oleh karena itu, perlu dilakukan sebuah kajian guna mengklasifikasikan kinerja asesor untuk memastikan bahwa asesor yang dipilih memiliki kemampuan yang tepat dan sesuai dengan kebutuhan BAN PAUD dan PNF Provinsi Sumatera Utara. Peneliti menerapkan metode klasifikasi naïve bayes dengan pengujian algoritmanya menggunakan Tools WEKA. Hasil akurasi dengan menggunakan Use Training Set, presentase untuk Correctly Classified Instance yaitu sebesar 97,0799% sementara persentase untuk Incorrectly Classified Instance adalah sebesar 2,9221%. Di mana dari 308 data penilaian kinerja asesor ada sebanyak 299 data asesor yang berhasil diklasifikasikan dengan benar dan sebanyak 9 data asesor tidak berhasil diklasifikasikan dengan benar.

## 1. PENDAHULUAN

Akreditasi adalah proses penilaian secara komprehensif terhadap kelayakan dan kinerja lembaga pendidikan atau program studi, yang dilakukan sebagai bagian dari sistem penjaminan mutu pendidikan [1]. Akreditasi satuan PAUD dan PNF merupakan proses penetapan formal yang dilakukan oleh BAN PAUD dan PNF untuk menentukan kualitas satuan PAUD dan PNF. Berdasarkan Peraturan Menteri Pendidikan dan Kebudayaan (Permendikbud) Nomor 13 Tahun 2018 tentang BAN S/M dan BAN PAUD dan PNF pada Pasal 1 Ayat (3) merupakan landasan awal dalam pelaksanaan akreditasi. Peraturan tersebut menjelaskan bahwa Badan Akreditasi Nasional Pendidikan Anak Usia Dini dan Pendidikan Nonformal yang selanjutnya disingkat BAN PAUD dan PNF adalah badan evaluasi mandiri yang menetapkan kelayakan satuan pendidikan anak usia dini dan pendidikan nonformal dengan mengacu pada Standar Nasional Pendidikan [2].

Dalam pelaksanaannya penilaian akreditasi dilakukan oleh asesor yang telah terlatih dan memiliki kompetensi yang sesuai. Asesor adalah tenaga profesional yang telah memenuhi persyaratan untuk diangkat dan ditugaskan oleh BAN untuk melakukan penilaian terhadap kelayakan Satuan Pendidikan sebagai bagian dari proses Akreditasi [3].

Asesor memiliki peran penting dalam mengevaluasi dan menilai kinerja lembaga pendidikan agar memenuhi standar yang ditetapkan. Penilaian

kinerja asesor menjadi hal yang krusial untuk memastikan bahwa lembaga pendidikan yang dinilai telah memenuhi kriteria yang dibutuhkan. Pentingnya memiliki asesor yang berkualitas dan kompeten untuk melakukan penilaian akreditasi terhadap lembaga Pendidikan Anak Usia Dini dan Pendidikan Nonformal di Provinsi Sumatera Utara. Namun, dalam kenyataannya tidak semua asesor memiliki kompetensi dan kualitas yang sama dalam melakukan penilaian akreditasi. Beberapa asesor mungkin lebih unggul dibandingkan dengan yang lain dalam hal pengalaman, keahlian, pengetahuan tentang pedagogi dan kurikulum. Karena itu, perlu dilakukan klasifikasi asesor yang objektif dan efektif untuk memastikan bahwa asesor yang dipilih memiliki kemampuan yang tepat dan sesuai dengan kebutuhan BAN PAUD dan PNF Provinsi Sumatera Utara.

Data Mining merupakan tahap dalam pengklasifikasian terhadap data dengan menghubungkan masing-masing pola pada setiap data set yang berukuran besar dengan jumlah data yang besar pula. Data Mining dapat juga didefinisikan menggali data dari banyaknya informasi yang akan dicari sehingga data yang perlu diketahui akan lebih mudah dicari dengan adanya sistem pola yang dibuat berdasarkan titik terdekat dengan informasi yang sering diperlukan [4]. Data Mining, sering juga disebut sebagai Knowledge Discovery in Database (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk

menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [5].

Berdasarkan hal tersebut, pada penelitian kali ini akan dilakukan menggunakan metode Naïve Bayes untuk dapat menganalisa dan mengklasifikasikan penilaian kinerja asesor dengan cara mengambil sample data penilaian kinerja asesor tahun 2021-2022 yang akan digunakan sebagai data latih dan menggunakan tool Weka untuk pengujian datanya.

Berdasarkan penelitian yang telah dilakukan oleh Viny Novika Sari, Lola Yorita Astri, Errissya Rasywir yaitu melakukan penelitian tentang penilaian karyawan yang terkadang dilakukan secara subjektif dan keterbatasan dalam mengontrol setiap karyawan yang bekerja pada PT. Pelita Wira Sejahtera. Oleh karena itu penulis melakukan analisis data mining pada data-data penilaian karyawan tersebut agar dapat mengetahui mana karyawan yang memiliki kinerja yang sangat baik, baik, cukup, dan kurang. Metode yang digunakan adalah metode klasifikasi Naive Bayes dengan persentasi akurasi terbesar diperoleh dengan menggunakan Use Training Set Correctly yaitu sebesar 95.302%, menggunakan 5-Fold Cross Validation Correctly sebesar 93,9597%, dan menggunakan 10-Fold Cross Validation sebesar 93.9597% [6].

Penelitian oleh Arif Senja Fitriani melakukan penelitian tentang partisipasi pemilu di Desa Jemirahan Kecamatan Jabon dengan menggunakan metode klasifikasi, yaitu algoritma Naïve Bayes. Data diperoleh dari Komisi Pemilihan Umum (KPU) Sidoarjo. Data tersebut di uji menggunakan metode klasifikasi Naive Bayes dengan Tools Weka dan Website dengan 6 variabel yang sudah ditentukan. Dataset di ambil sebanyak 300 data dibagi 2 yaitu sebanyak 65% dari 195 data Training dan sebanyak 35% dari 105 data Testing. Hasil prediksi partisipasi pemilu dari dataset yang diambil sebanyak 300 data dibagi 2 yaitu sebanyak 65% dari 195 data Training dan sebanyak 35% dari 105 data Testing. Hasil prediksi berdasarkan set atribut kehadiran, untuk kehadiran dengan 105 data diperoleh nilai 97% prediksi kebenarannya dan diperoleh nilai 3% prediksi kesalahannya [7].

## 2. METODE

### 2.1. Metode Pengumpulan Data

Penelitian ini dilakukan di Balai Penjaminan Mutu Pendidikan (BPMP) Provinsi Sumatera Utara yang berlokasi di Jl. Bunga Raya No. 96, Asam Kumbang Kec. Medan Selayang, Kota Medan sebagai objek penelitian. Dalam mencari data peneliti menggunakan dua teknik yaitu wawancara langsung dengan narasumber untuk mendapatkan data yang nantinya akan diolah menggunakan metode *Naïve Bayes* dan studi kepustakaan untuk memberikan landasan teori yang kuat melalui buku-buku maupun jurnal yang berhubungan dengan penelitian ini.

Dalam penelitian ini, data yang digunakan adalah data penilaian kinerja asesor pada BAN PAUD dan PNF Provinsi Sumatera Utara Tahun 2021-2022 berjumlah 308 data. Data ini terdiri dari beberapa atribut yaitu jenis kelamin, rumpun, nilai penguasaan substansi, nilai penguasaan TIK/Sispena, nilai kemampuan melaksanakan tugas, nilai integritas dan grade dari asesor yang selanjutnya akan dilakukan proses pengolahan data mining antara lain yaitu proses *Data Selection*, *Preprocessing/Cleaning*, *Transformation*, dan *Data mining* yang kemudian akan dilakukan uji coba dengan menggunakan *tools* WEKA.

### 2.2. Metode Naïve Bayes

Pada penelitian ini algoritma yang digunakan yaitu algoritma Naïve Bayes. Naïve Bayes merupakan suatu algoritma yang dapat bertugas dalam melakukan pengklasifikasian suatu variable dalam kumpulan basis data sesuai dengan perhitungan yang mengacu pada bidang ilmu matematika yaitu menggunakan metode perhitungan probabilitas dan statistik [8]. Salah satu metode yang mengasumsikan bahwa nilai antar variabel saling bebas (independen) yang tidak terpengaruh oleh variabel kelas [9]. Adapun kelebihan dari penggunaan metode naïve bayes yaitu hanya memerlukan jumlah data pelatihan (training data) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian [10]. Persamaan dari algoritma yang digunakan adalah sebagai berikut [11] :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

## 3. HASIL DAN PEMBAHASAN

### 3.1. Perhitungan Metode Naïve Bayes

Pada penelitian ini algoritma yang digunakan adalah algoritma Naïve Bayes. Adapun langkah-langkah penerapan metode perhitungan menggunakan Naïve Bayes adalah sebagai berikut :

#### 3.1.1. Baca data latih (training)

Untuk menentukan data yang nantinya akan dianalisis dengan menggunakan metode Naïve Bayes maka langkah pertama yang dilakukan adalah membaca data latih. Data yang digunakan adalah data penilaian kinerja asesor dari tahun 2021-2022 sebanyak 308 data. Aspek yang dinilai meliputi: penilaian terhadap penguasaan substansi, penguasaan TIK/Sispena, kemampuan dalam melaksanakan tugas dan integritas dari asesor. Penilaian dilakukan dengan menggunakan skala likert dengan rentang nilai 1-5 yaitu sangat tidak baik, tidak baik, cukup, baik dan sangat baik. Nilai grade dihasilkan dari penjumlahan nilai setiap kompetensi dibagi dengan banyaknya kategori sehingga didapatkan grade sangat baik, baik dan cukup. Adapun data latih yang digunakan dapat dilihat pada tabel 1.

Tabel 1. Tabel Data Latih (*Training*)

No	Jenis Kelamin	Rumpun (PAUD/ PKBM)	Nilai Penguasaan Substansi	Nilai Penguasaan TIK/ SisPena	Nilai Kemampuan Melaksanakan Tugas	Nilai Integritas	Grade
1	P	PAUD	Cukup	Baik	Cukup	Cukup	Cukup
2	L	PAUD	Baik	Baik	Cukup	Baik	Baik
3	L	PAUD	Cukup	Baik	Cukup	Baik	Cukup
4	P	PAUD	Baik	Baik	Baik	Baik	Baik
5	P	PAUD	Baik	Baik	Baik	Tidak Baik	Cukup
6	L	PKBM	Baik	Cukup	Cukup	Cukup	Cukup
7	L	PKBM	Baik	Baik	Baik	Sangat Baik	Sangat Baik
8	L	PKBM	Baik	Baik	Baik	Baik	Baik
9	P	PKBM	Baik	Sangat Baik	Baik	Baik	Sangat Baik
10	P	PAUD	Sangat Baik	Baik	Sangat Baik	Baik	Sangat Baik
11	P	PAUD	Sangat Baik	Baik	Baik	Baik	Sangat Baik
12	P	PKBM	Cukup	Tidak Baik	Cukup	Cukup	Cukup
13	L	PAUD	Cukup	Baik	Baik	Cukup	Cukup
14	L	PKBM	Baik	Sangat Baik	Baik	Baik	Sangat Baik
15	P	PAUD	Sangat Baik	Baik	Baik	Sangat Baik	Sangat Baik
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
308	P	PAUD	Cukup	Cukup	Cukup	Cukup	Cukup

3.1.2. Kriteria dan Perhitungan Probabilitas

Adapun nilai probabilitas dari masing-masing kriteria yang dihitung berdasarkan dari data latih pada tabel 1, dengan menerapkan rumus persamaan 1 sebagai berikut :

a. Probabilitas Kriteria Nilai Penguasaan Substansi

Berdasarkan data latih penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor, dimana dari 308 data terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai cukup dengan grade sangat baik, 36 data asesor memiliki nilai baik dengan grade sangat baik, dan 44 data asesor memiliki nilai sangat baik dengan grade sangat baik. Lalu terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade baik, 0 data asesor memiliki nilai tidak baik dengan grade baik, 10 data asesor memiliki nilai cukup dengan grade baik, 68 data asesor memiliki nilai baik dengan grade baik, dan 2 data asesor memiliki nilai sangat baik dengan grade baik. Selanjutnya terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade cukup, 4 data asesor memiliki nilai tidak baik dengan grade cukup, 127 data asesor memiliki nilai cukup dengan grade cukup, 17 data asesor memiliki nilai baik dengan grade cukup, dan 0 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas kriteria terhadap nilai penguasaan substansi dapat dilihat pada tabel 2.

Tabel 2. Probabilitas Nilai Penguasaan Substansi

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	Baik	Cukup	Sangat baik	Baik	Cukup
Sangat Tidak Baik	0	0	0	$0/80 = 0,00$	$0/80 = 0,00$	$0/148 = 0,00$

Tidak Baik	0	0	4	$0/80 = 0,00$	$0/80 = 0,00$	$4/148 = 0,03$
Cukup	0	10	127	$0/80 = 0,00$	$10/80 = 0,13$	$127/148 = 0,86$
Baik	36	68	17	$36/80 = 0,45$	$68/80 = 0,85$	$17/148 = 0,11$
Sangat Baik	44	2	0	$44/80 = 0,55$	$2/80 = 0,03$	$0/148 = 0,00$
Jumlah	80	80	148	1	1	1

b. Probabilitas Kriteria Nilai Penguasaan TIK/Sispna

Pada kriteria nilai penguasaan TIK/Sispna data berjumlah 308 data asesor, terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai cukup dengan grade sangat baik, 45 data asesor memiliki nilai baik dengan grade sangat baik, dan 35 data asesor memiliki nilai sangat baik dengan grade sangat baik. Lalu terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade baik, 0 data asesor memiliki nilai tidak baik dengan grade baik, 14 data asesor memiliki nilai cukup dengan grade baik, 64 data asesor memiliki nilai baik dengan grade baik, dan 2 data asesor memiliki nilai sangat baik dengan grade baik. Selanjutnya terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade cukup, 6 data asesor memiliki nilai tidak baik dengan grade cukup, 118 data asesor memiliki nilai cukup dengan grade cukup, 24 data asesor memiliki nilai baik dengan grade cukup, dan 0 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas kriteria terhadap nilai penguasaan substansi dapat dilihat pada tabel 3.

Tabel 3. Probabilitas Nilai Penguasaan TIK/Sispna

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Sangat Tidak Baik	0	0	0	$0/80 = 0,00$	$0/80 = 0,00$	$0/148 = 0,00$
Tidak Baik	0	0	6	$0/80 = 0,00$	$0/80 = 0,00$	$6/148 = 0,04$
Cukup	0	14	118	$0/80 = 0,00$	$14/80 = 0,18$	$118/148 = 0,80$
Baik	45	64	24	$45/80 = 0,56$	$64/80 = 0,80$	$24/148 = 0,16$
Sangat Baik	35	2	0	$35/80 = 0,44$	$2/80 = 0,03$	$0/148 = 0,00$
Jumlah	80	80	148	1	1	1

c. Probabilitas Kriteria Nilai Kemampuan Melaksanakan Tugas

Pada kriteria nilai kemampuan melaksanakan tugas data berjumlah 308 data asesor, terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai tidak baik dengan grade sangat baik, 1 data asesor memiliki nilai cukup dengan grade sangat baik, 40 data asesor memiliki nilai baik dengan grade sangat baik, dan 39 data asesor memiliki nilai sangat baik dengan grade sangat baik. Lalu terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade baik, 0 data asesor memiliki nilai tidak baik dengan grade baik, 4 data asesor memiliki nilai cukup dengan grade baik, 75 data asesor memiliki nilai baik dengan grade baik, dan 1 data asesor memiliki nilai sangat baik dengan grade baik. Selanjutnya terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade cukup, 6 data asesor memiliki nilai tidak baik dengan grade cukup, 131 data asesor memiliki nilai cukup dengan grade cukup, 11 data asesor memiliki nilai baik dengan grade cukup, dan 0 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas kriteria terhadap nilai penguasaan substansi dapat dilihat pada tabel 4.

Tabel 4. Probabilitas Nilai Kemampuan Melaksanakan Tugas

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Sangat Tidak Baik	0	0	0	$0/80 = 0,00$	$0/80 = 0,00$	$0/148 = 0,00$
Tidak Baik	0	0	6	$0/80 = 0,00$	$0/80 = 0,00$	$6/148 = 0,04$
Cukup	1	4	131	$1/80 = 0,01$	$4/80 = 0,05$	$131/148 = 0,89$
Baik	40	75	11	$40/80 = 0,50$	$75/80 = 0,94$	$11/148 = 0,07$
Sangat Baik	39	1	0	$39/80 = 0,49$	$1/80 = 0,01$	$0/148 = 0,00$
Jumlah	80	80	148	1	1	1

d. Probabilitas Kriteria Nilai Integritas

Pada kriteria nilai integritas data berjumlah 308 data asesor, terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade sangat baik, 0 data asesor memiliki nilai tidak baik dengan grade sangat baik, 1 data asesor memiliki nilai cukup dengan grade sangat baik, 19 data asesor memiliki nilai baik dengan grade sangat baik, dan 60 data asesor memiliki nilai sangat baik dengan grade sangat baik. Lalu terdapat 0 data asesor memiliki nilai sangat tidak baik dengan grade baik, 0 data asesor memiliki nilai tidak baik dengan grade baik, 13 data asesor memiliki nilai cukup dengan grade baik, 64 data asesor memiliki nilai baik dengan grade baik, dan 3 data asesor memiliki nilai sangat baik dengan grade baik. Selanjutnya terdapat 1 data asesor memiliki nilai sangat tidak baik dengan grade cukup, 7 data asesor memiliki nilai tidak baik dengan grade cukup, 122 data asesor memiliki nilai cukup dengan grade cukup, 18 data asesor memiliki nilai baik dengan grade cukup, dan 0 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas kriteria terhadap nilai penguasaan substansi dapat dilihat pada tabel 5.

Tabel 5. Probabilitas Nilai Integritas

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Sangat Tidak Baik	0	0	1	$0/80 = 0,00$	$0/80 = 0,00$	$1/148 = 0,01$
Tidak Baik	0	0	7	$0/80 = 0,00$	$0/80 = 0,00$	$7/148 = 0,05$
Cukup	1	13	122	$1/80 = 0,01$	$13/80 = 0,16$	$122/148 = 0,82$
Baik	19	64	18	$19/80 = 0,24$	$64/80 = 0,80$	$18/148 = 0,12$
Sangat Baik	60	3	0	$60/80 = 0,75$	$3/80 = 0,04$	$0/148 = 0,00$
Jumlah	80	80	148	1	1	1

e. Probabilitas Kriteria Jenis Kelamin

Diketahui dari 308 data asesor, terdapat 36 data asesor dengan jenis kelamin laki-laki memiliki grade sangat baik, 21 data asesor dengan jenis kelamin laki-laki memiliki grade baik, dan 43 data asesor dengan jenis kelamin laki-laki memiliki grade cukup. Lalu terdapat 44 data asesor dengan jenis kelamin perempuan memiliki grade sangat baik, 59 data asesor dengan jenis kelamin perempuan memiliki grade baik, dan 105 data asesor dengan jenis kelamin perempuan memiliki grade cukup. Probabilitas kriteria jenis kelamin dapat dilihat pada tabel 6.

Tabel 6. Probabilitas Jenis Kelamin

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup

Laki-laki	36	21	43	36/80 = 0,45	21/80 = 0,26	43/148 = 0,29
Perempuan	44	59	105	44/80 = 0,55	59/80 = 0,74	105/148 = 0,71
Jumlah	80	80	148	1	1	1

f. Probabilitas Kriteria Rumpun

Diketahui dari 308 data asesor, terdapat 50 data asesor dari rumpun PAUD dengan grade sangat baik, 72 data asesor dari rumpun PAUD dengan grade baik, dan 129 data asesor dari rumpun PAUD dengan grade cukup. Selain itu juga terdapat 30 data asesor dari rumpun PKBM dengan grade sangat baik, 8 data asesor dari rumpun PKBM dengan grade baik, dan 19 data asesor dari rumpun PKBM dengan grade cukup. Probabilitas kriteria rumpun dapat dilihat pada tabel 7.

Tabel 7. Probabilitas Rumpun

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
PAUD	50	72	129	50/80 = 0,63	72/80 = 0,90	129/148 = 0,87
PKBM	30	8	19	30/80 = 0,38	8/80 = 0,10	19/148 = 0,13
Jumlah	80	80	148	1	1	1

Pada tabel probabilitas masing-masing kriteria di atas terlihat bahwa masih banyak kolom yang bernilai 0, dalam *Naïve Bayes* hal ini menjadi anomali karena tidak boleh terdapat nilai 0 dalam probabilitas karena nilai 0 menyatakan “tidak mungkin terjadi”. *Anomaly* tersebut kemudian dapat dikoreksi dengan menggunakan teknik perhitungan *Laplacian Correction*. *Laplacian Correction* atau *Laplacian Smoothie* adalah teknik perhitungan untuk menyiasati supaya probabilitas pada perhitungan *Naïve Bayes* tidak bernilai 0, dikarenakan tidak adanya data untuk kategori tertentu dalam sebuah *class* [12].

**Laplace Correction**

*Laplace Correction (Laplacian Estimator)* atau *additive smoothing* adalah suatu cara untuk mengurangi kemungkinan kegagalan klasifikasi, karena bisa mencegah hasil klasifikasi yang probabilitasnya bernilai 0 (nol). Dari sekian banyaknya data pada training set, untuk setiap perhitungan atribut ditambah nilai 1 (satu) dan tidak akan membuat perbedaan yang berarti pada estimasi probabilitas sehingga bisa menghindari kasus dengan nilai probabilitas 0 (nol) [13]. *Laplace Correction* ini dapat dinyatakan dalam persamaan berikut ini [14] :

$$p_i = \frac{m_i + 1}{n + k}$$

Keterangan :

$p_i$  : probabilitas dari atribut  $m_i$

$m_i$  : jumlah sampel dalam kelas dari atribut  $m_i$   
 $k$  : jumlah kelas dari atribut  $m_i$   
 $n$  : jumlah sampel

Adapun perhitungan dari tiap probabilitas yang telah dihitung dengan menggunakan rumus persamaan *laplace correction* adalah sebagai berikut :

a. Kriteria Nilai Penguasaan Substansi dengan *Laplace Correction*

Berdasarkan data latihan penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan *laplace correction* maka data bertambah menjadi 320 data yaitu terdapat 4 (empat) atribut : Tidak Baik, Cukup, Baik, dan Sangat Baik dengan 3 (tiga) kelas : Sangat Baik, Baik dan Cukup, dimana cara kerjanya dengan menambahkan 1 tupel atau *record* pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 1 data asesor memiliki nilai tidak baik dengan grade sangat baik, 1 data asesor memiliki nilai cukup dengan grade sangat baik, 37 data asesor memiliki nilai baik dengan grade sangat baik, 45 data asesor memiliki nilai sangat baik dengan grade sangat baik, 1 data asesor memiliki nilai tidak baik dengan grade baik, 11 data asesor memiliki nilai cukup dengan grade baik, 69 data asesor memiliki nilai baik dengan grade baik, 3 data asesor memiliki nilai sangat baik dengan grade baik, 5 data asesor memiliki nilai tidak baik dengan grade cukup, 128 data asesor memiliki nilai cukup dengan grade cukup, 18 data asesor memiliki nilai baik dengan grade cukup, 1 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas nilai penguasaan substansi dapat dilihat pada tabel 8.

Tabel 8. Probabilitas Nilai Penguasaan Substansi dengan *Laplace Correction*

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Tidak Baik	1	1	5	1/84 = 0,01	1/84 = 0,01	5/152 = 0,03
Cukup	1	11	128	1/84 = 0,01	11/84 = 0,13	128/152 = 0,84
Baik	37	69	18	37/84 = 0,44	69/84 = 0,82	18/152 = 0,12
Sangat Baik	45	3	1	45/84 = 0,54	3/84 = 0,04	1/152 = 0,01
Jumlah	84	84	152	1	1	1

b. Probabilitas Kriteria Nilai Penguasaan TIK/Sispena dengan *Laplace Correction*

Berdasarkan data latihan penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan *laplace correction* maka data bertambah menjadi 320 data yaitu terdapat 4 (empat) atribut: Tidak Baik, Cukup, Baik, dan Sangat Baik dengan 3 (tiga) kelas : Sangat Baik, Baik dan Cukup, dimana cara kerjanya

dengan menambahkan 1 tupel atau record pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 1 data asesor memiliki nilai tidak baik dengan grade sangat baik, 1 data asesor memiliki nilai cukup dengan grade sangat baik, 46 data asesor memiliki nilai baik dengan grade sangat baik, 36 data asesor memiliki nilai sangat baik dengan grade sangat baik, 1 data asesor memiliki nilai tidak baik dengan grade baik, 15 data asesor memiliki nilai cukup dengan grade baik, 65 data asesor memiliki nilai baik dengan grade baik, 3 data asesor memiliki nilai sangat baik dengan grade baik, 7 data asesor memiliki nilai tidak baik dengan grade cukup, 119 data asesor memiliki nilai cukup dengan grade cukup, 25 data asesor memiliki nilai baik dengan grade cukup, 1 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas nilai penguasaan TIK/Sispna dapat dilihat pada tabel 9.

Tabel 9. Probabilitas Nilai Penguasaan TIK/Sispna dengan Laplace Correction

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Tidak Baik	1	1	7	$1/84 = 0,01$	$1/84 = 0,01$	$7/152 = 0,05$
Cukup	1	15	119	$1/84 = 0,01$	$15/84 = 0,18$	$119/152 = 0,78$
Baik	46	65	25	$46/84 = 0,55$	$65/84 = 0,77$	$25/152 = 0,16$
Sangat Baik	36	3	1	$36/84 = 0,43$	$3/84 = 0,04$	$1/152 = 0,01$
Jumlah	84	84	152	1	1	1

c. Probabilitas Kriteria Nilai Kemampuan Melaksanakan Tugas dengan Laplace Correction

Berdasarkan data latih penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan laplace correction maka data bertambah menjadi 320 data yaitu terdapat 4 (empat) atribut: Tidak Baik, Cukup, Baik, dan Sangat Baik dengan 3 (tiga) kelas : Sangat Baik, Baik dan Cukup, dimana cara kerjanya dengan menambahkan 1 tupel atau record pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 1 data asesor memiliki nilai tidak baik dengan grade sangat baik, 2 data asesor memiliki nilai cukup dengan grade sangat baik, 41 data asesor memiliki nilai baik dengan grade sangat baik, 40 data asesor memiliki nilai sangat baik dengan grade sangat baik, 1 data asesor memiliki nilai tidak baik dengan grade baik, 5 data asesor memiliki nilai cukup dengan grade baik, 76 data asesor memiliki nilai baik dengan grade baik, 2 data asesor memiliki nilai sangat baik dengan grade baik, 7 data asesor memiliki nilai tidak baik dengan grade cukup, 132 data asesor

memiliki nilai cukup dengan grade cukup, 12 data asesor memiliki nilai baik dengan grade cukup, 1 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas nilai kemampuan melaksanakan tugas dapat dilihat pada tabel 10.

Tabel 10. Probabilitas Nilai Kemampuan Melaksanakan Tugas dengan Laplace Correction

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Tidak Baik	1	1	7	$1/84 = 0,01$	$1/84 = 0,01$	$7/152 = 0,05$
Cukup	2	5	132	$2/84 = 0,02$	$5/84 = 0,06$	$132/152 = 0,87$
Baik	41	76	12	$41/84 = 0,49$	$76/84 = 0,90$	$12/152 = 0,08$
Sangat Baik	40	2	1	$40/84 = 0,48$	$2/84 = 0,02$	$1/152 = 0,01$
Jumlah	84	84	152	1	1	1

d. Probabilitas Kriteria Nilai Integritas dengan Laplace Correction

Berdasarkan data latih penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan laplace correction maka data bertambah menjadi 323 data yaitu terdapat 5 (lima) atribut : Sangat Tidak Baik, Tidak Baik, Cukup, Baik, dan Sangat Baik dengan 3 (tiga) kelas: Sangat Baik, Baik dan Cukup, dimana cara kerjanya dengan menambahkan 1 tupel atau record pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 1 data asesor memiliki nilai sangat tidak baik dengan grade sangat baik, 1 data asesor memiliki nilai tidak baik dengan grade sangat baik, 2 data asesor memiliki nilai cukup dengan grade sangat baik, 20 data asesor memiliki nilai baik dengan grade sangat baik, 61 data asesor memiliki nilai sangat baik dengan grade sangat baik, 1 data asesor memiliki nilai sangat tidak baik dengan grade baik, 1 data asesor memiliki nilai tidak baik dengan grade baik, 14 data asesor memiliki nilai cukup dengan grade baik, 65 data asesor memiliki nilai baik dengan grade baik, 4 data asesor memiliki nilai sangat baik dengan grade baik, 2 data asesor memiliki nilai sangat tidak baik dengan grade cukup, 8 data asesor memiliki nilai tidak baik dengan grade cukup, 123 data asesor memiliki nilai cukup dengan grade cukup, 19 data asesor memiliki nilai baik dengan grade cukup, 1 data asesor memiliki nilai sangat baik dengan grade cukup. Probabilitas nilai integritas dapat dilihat pada tabel 11.

Tabel 11. Probabilitas Nilai Integritas dengan Laplace Correction

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	baik	cukup	Sangat baik	baik	cukup
Sangat Tidak Baik	1	1	2	$1/85 = 0,01$	$1/85 = 0,01$	$2/153 = 0,01$

Tidak Baik	1	1	8	$1/85 = 0,01$	$1/85 = 0,01$	$8/153 = 0,05$
Cukup	2	14	123	$2/85 = 0,02$	$14/85 = 0,17$	$123/153 = 0,81$
Baik	20	65	19	$20/85 = 0,24$	$65/85 = 0,77$	$19/153 = 0,13$
Sangat Baik	61	4	1	$61/85 = 0,73$	$4/85 = 0,05$	$1/153 = 0,01$
Jumlah	85	85	153	1	1	1

e. Probabilitas Kriteria Jenis Kelamin dengan Laplace Correction

Berdasarkan data latih penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan laplace correction maka data bertambah menjadi 314 data yaitu terdapat 2 (dua) atribut : Laki-laki dan Perempuan dengan 3 (tiga) kelas : Sangat Baik, Baik dan Cukup, dimana cara kerjanya dengan menambahkan 1 tupel atau record pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 37 data asesor dengan jenis kelamin laki-laki dan grade sangat baik, 45 data asesor dengan jenis kelamin perempuan dan grade sangat baik, 22 data asesor dengan jenis kelamin laki-laki dan grade baik, 60 data asesor dengan jenis kelamin perempuan dan grade baik, 44 data asesor dengan jenis kelamin laki-laki dan grade cukup 106 data asesor dengan jenis kelamin perempuan dan grade cukup. Probabilitas kriteria jenis kelamin dapat dilihat pada tabel 12.

Tabel 12. Probabilitas Jenis Kelamin

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	Baik	Cukup	Sangat baik	Baik	Cukup
Laki-laki	37	22	44	$37/82 = 0,45$	$22/82 = 0,27$	$44/150 = 0,29$
Perempuan	45	60	106	$45/82 = 0,55$	$60/82 = 0,73$	$106/150 = 0,71$
Jumlah	82	82	150	1	1	1

f. Probabilitas Kriteria Rumpun dengan Laplace Correction

Berdasarkan data latih penilaian kinerja asesor pada tabel 1 diketahui data berjumlah 308 data asesor. Setelah melakukan laplace correction maka data bertambah menjadi 314 data yaitu terdapat 2 (dua) atribut : PAUD dan PKBM dengan 3 (tiga) kelas : Sangat Baik, Baik dan Cukup, dimana cara kerjanya dengan menambahkan 1 tupel atau record pura-pura pada semua atribut untuk menghindari nilai probabilitas 0 (nol).

Adapun datanya yaitu terdapat 51 data asesor dengan rumpun PAUD memiliki grade sangat baik, 31 data asesor dengan rumpun PKBM memiliki grade sangat baik, 73 data asesor dengan rumpun PAUD memiliki grade baik, 9 data asesor dengan rumpun PKBM memiliki grade baik, 130 data asesor dengan rumpun PAUD memiliki grade cukup, 20 data asesor

dengan rumpun PKBM memiliki grade cukup. Probabilitas kriteria rumpun dapat dilihat pada tabel 13.

Tabel 13. Probabilitas Rumpun

Nilai Penguasaan Substansi	Grade			Probabilitas		
	Sangat baik	Baik	Cukup	Sangat baik	Baik	Cukup
PAUD	51	73	130	$51/82 = 0,62$	$73/82 = 0,89$	$130/150 = 0,87$
PKBM	31	9	20	$31/82 = 0,38$	$9/82 = 0,11$	$20/150 = 0,13$
Jumlah	82	82	150	1	1	1

Pengujian Metode Naïve Bayes

Data Testing (Data Uji)

Berikut adalah perhitungan manual yang dilakukan dengan menggunakan Microsoft Excel guna mempercepat untuk mendapatkan hasil perhitungannya dengan metode Naïve Bayes.

No	Jenis Kelamin	Rumpun (PAUD/PKB M)	Nilai Penguasaan Substansi	Nilai Penguasaan TIK/SisPena	Nilai Kemampuan Melaksanakan Tugas	Nilai Integritas	Grade
1	L	PAUD	Sangat Baik	Sangat Baik	Cukup	Baik	Sangat Baik

Adapun langkah-langkah perhitungan probabilitasnya adalah sebagai berikut :

- Menghitung jumlah kelas/label
  - Kelas "Sangat Baik" =  $84/320 = 0,26$
  - Kelas "Baik" =  $84/320 = 0,26$
  - Kelas "Cukup" =  $152/320 = 0,48$
- Menghitung jumlah kriteria yang sama dengan kelas/label yang sama
  - $P(\text{Jenis\_Kelamin} = \text{"Laki-laki"} \mid \text{Grade} = \text{"Sangat Baik"}) = 37/82 = 0,45$
  - $P(\text{Jenis\_Kelamin} = \text{"Laki-Laki"} \mid \text{Grade} = \text{"Baik"}) = 22/82 = 0,27$
  - $P(\text{Jenis\_Kelamin} = \text{"Laki-Laki"} \mid \text{Grade} = \text{"Cukup"}) = 44/150 = 0,29$
  - $P(\text{Rumpun} = \text{"PAUD"} \mid \text{Grade} = \text{"Sangat Baik"}) = 51/82 = 0,62$
  - $P(\text{Rumpun} = \text{"PAUD"} \mid \text{Grade} = \text{"Baik"}) = 73/82 = 0,89$
  - $P(\text{Rumpun} = \text{"PAUD"} \mid \text{Grade} = \text{"Cukup"}) = 130/150 = 0,87$
  - $P(\text{Nilai\_Penguasaan\_Substansi} = \text{"Sangat Baik"} \mid \text{Grade} = \text{"Sangat Baik"}) = 45/84 = 0,54$
  - $P(\text{Nilai\_Penguasaan\_Substansi} = \text{"Sangat Baik"} \mid \text{Grade} = \text{"Baik"}) = 3/84 = 0,04$
  - $P(\text{Nilai\_Penguasaan\_Substansi} = \text{"Sangat Baik"} \mid \text{Grade} = \text{"Cukup"}) = 1/152 = 0,01$
  - $P(\text{Nilai\_Penguasaan\_TIK/Sispena} = \text{"Sangat Baik"} \mid \text{Grade} = \text{"Sangat Baik"}) = 36/84 = 0,43$

$$P(\text{Nilai\_Penguasaan\_TIK/Sispna} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Baik”}) = 3/84 = 0,04$$

$$P(\text{Nilai\_Penguasaan\_TIK/Sispna} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Cukup”}) = 1/152 = 0,01$$

$$P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Sangat Baik”}) = 2/84 = 0,02$$

$$P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Baik”}) = 5/84 = 0,06$$

$$P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Cukup”}) = 132/152 = 0,87$$

$$P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Sangat Baik”}) = 20/85 = 0,24$$

$$P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Baik”}) = 65/85 = 0,77$$

$$P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Cukup”}) = 19/153 = 0,13$$

c. Mengalikan semua hasil variabel setiap klasifikasi

$$P(\text{Jenis\_Kelamin} = \text{“Laki-laki”} \mid \text{Grade} = \text{“Sangat Baik”}) \times P(\text{Rumpun} = \text{“PAUD”} \mid \text{Grade} = \text{“Sangat Baik”}) \times P(\text{Nilai\_Penguasaan\_Substansi} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Sangat Baik”}) \times P(\text{Nilai\_Penguasaan\_TIK/Sispna} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Sangat Baik”}) \times P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Sangat Baik”}) \times P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Sangat Baik”})$$

$$= 37/82 \times 51/82 \times 45/84 \times 36/84 \times 2/84 \times 20/85$$

$$= 0,45 \times 0,62 \times 0,54 \times 0,43 \times 0,02 \times 0,24$$

$$= 0,000365$$

$$P(\text{Jenis\_Kelamin} = \text{“Laki-laki”} \mid \text{Grade} = \text{“Baik”}) \times P(\text{Rumpun} = \text{“PAUD”} \mid \text{Grade} = \text{“Baik”}) \times P(\text{Nilai\_Penguasaan\_Substansi} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Baik”}) \times P(\text{Nilai\_Penguasaan\_TIK/Sispna} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Baik”}) \times P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Baik”}) \times P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Baik”})$$

$$= 22/82 \times 73/82 \times 3/84 \times 3/84 \times 5/84 \times 65/85$$

$$= 0,27 \times 0,89 \times 0,04 \times 0,04 \times 0,06 \times 0,077$$

$$= 0,000014$$

$$P(\text{Jenis\_Kelamin} = \text{“Laki-laki”} \mid \text{Grade} = \text{“Cukup”}) \times P(\text{Rumpun} = \text{“PAUD”} \mid \text{Grade} = \text{“Cukup”}) \times P(\text{Nilai\_Penguasaan\_Substansi} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Cukup”}) \times P(\text{Nilai\_Penguasaan\_TIK/Sispna} = \text{“Sangat Baik”} \mid \text{Grade} = \text{“Cukup”}) \times P(\text{Nilai\_Kemampuan\_Melaksanakan\_Tugas} = \text{“Cukup”} \mid \text{Grade} = \text{“Cukup”}) \times P(\text{Nilai\_Integritas} = \text{“Baik”} \mid \text{Grade} = \text{“Cukup”})$$

$$= 44/150 \times 130/150 \times 1/152 \times 1/152 \times 132/152 \times 19/153$$

$$= 0,29 \times 0,87 \times 0,01 \times 0,01 \times 0,87 \times 0,013$$

$$= 0,000001$$

d. Membandingkan hasil pada setiap class yang di klasifikasikan

Maka dari perbandingan hasil penjumlahan diatas, bahwa nilai probabilitas tertinggi ada pada class (P | Grade = “Sangat Baik”) dengan nilai 0,000365 sehingga dapat disimpulkan bahwa pengujian pada data asesor ke-272 yang diuji yaitu memperoleh Grade “Sangat Baik”.

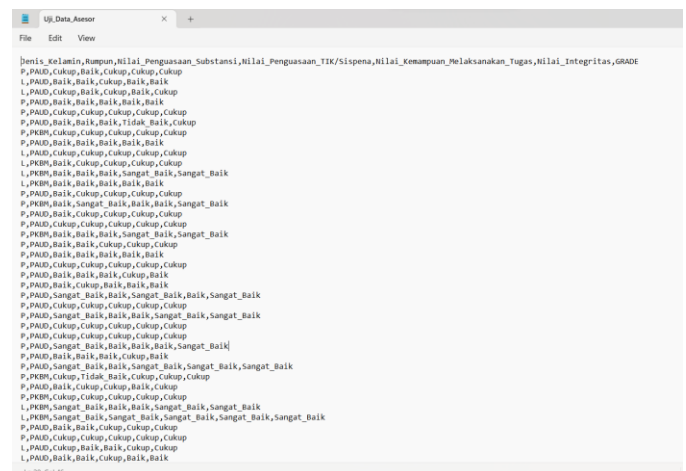
### 3.2. Implementasi Menggunakan Tools WEKA

#### 3.8.5

*Weka* mempunyai kelebihan seperti mempunyai banyak algoritma *data mining* dan *machine learning*, kemudahan dalam penggunaannya, selalu *up-to-date* dengan algoritma-algoritma yang baru. Semua teknik *Weka* adalah didasarkan pada asumsi bahwa data tersedia sebagai *flat file* tunggal atau hubungan, dimana setiap titik data digambarkan oleh sejumlah atribut (biasanya, atribut numerik atau nominal, tetapi beberapa jenis atribut lain juga didukung) [15].

Adapun tahapan menggunakan *tools WEKA* 3.8.5 adalah sebagai berikut :

- 1) Buka aplikasi *Microsoft Excel* lalu kemudian *openfile* dengan nama “Uji\_Data\_Aesor”.
- 2) Data dari hasil Transformasi di *Excel* kemudian di *block* disalin dan dimasukkan kedalam *notepad* untuk tanda *space* harus diganti dengan tanda “\_” dan *space* antar atribut diubah menjadi tanda “,” agar data tersebut dapat dibaca oleh format *.csv*. Bentuk transformasi data dari format *.xlsx* ke format *.csv* dapat dilihat pada Gambar 2 berikut.



Gambar 2. Transformasi data .CSV.

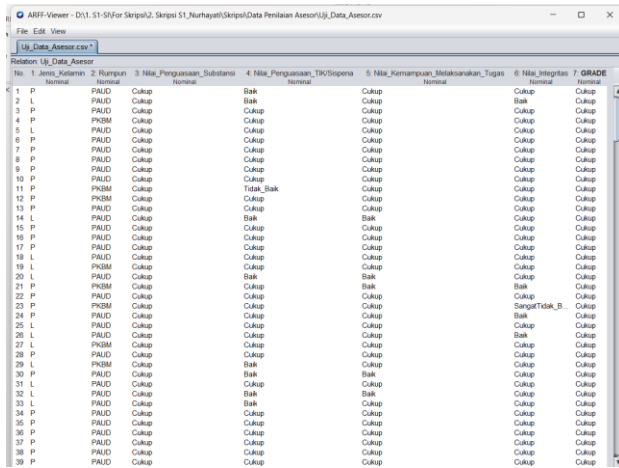
- 3) Selanjutnya buka *tools WEKA* 3.8.5. Berikut tampilan *windows* untuk *Weka GUI Chooser* dapat dilihat pada Gambar 3 berikut.



Gambar 3. Tampilan Menu Utama WEKA 3.8.5.

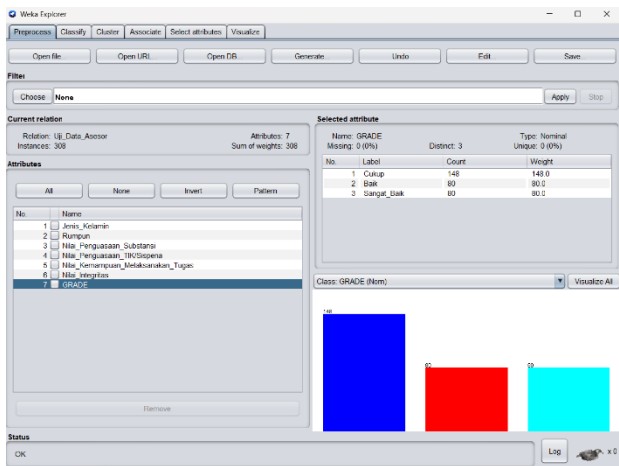


4) Buka “tools” dan buka *ArffViewer* open data hasil transformasi kedalam bentuk .csv kemudian kemudian *save as* kedalam bentuk *Arff*. Perubahan data dalam bentuk *Arff* dapat dilihat pada Gambar 4 berikut.



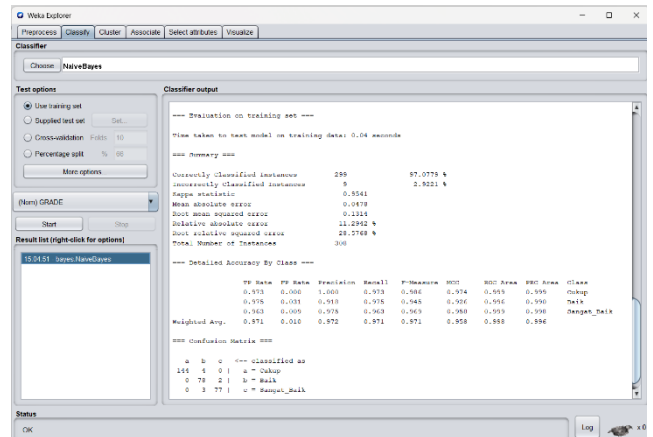
Gambar 4. Transformasi data *Arff*.

5) Kemudian buka file “Uji\_Data\_Asesor.arff”. Data *preprocessing* menampilkan masing-masing atribut yang digunakan. Setelah menyimpan data “Uji\_Data\_Asesor.arff”, lalu kembali ke menu utama dan pilih menu *Explorer*. Data *preprocessing* untuk file Uji Data Asesor dapat dilihat pada Gambar 5.



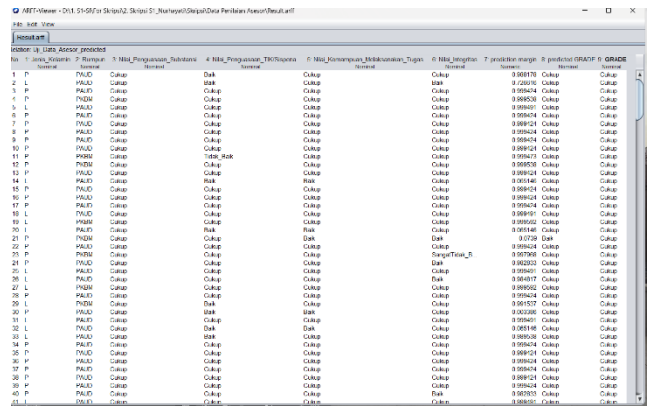
Gambar 5. Tampilan atribut *preprocessing* Uji Data Asesor

6) Kemudian pilih *Classify* dan pilih *Choose* lalu pilih *Bayes* kemudian *Naive Bayes*. Tampilan implementasi algoritma *Naive Bayes* untuk nilai akurasi dapat dilihat pada Gambar 6.



Gambar 6. Tampilan Implementasi Algoritma *Naive Bayes*.

7) Hasil dari pengujian metode untuk menemukan hasil dapat dilihat pada Gambar 7 berikut.



Gambar 7. Hasil Data yang telah di Uji

#### 4. KESIMPULAN

Berdasarkan pembahasan dari penelitian yang dilakukan, maka dapat diambil beberapa kesimpulan sebagai berikut :

- 1) Berdasarkan data penilaian kinerja asesor yang diperoleh, proses data mining membantu dalam penerapan metode Naïve Bayes dan teknik Laplacian Correction dengan mengolah data set sebanyak 308 data sehingga didapatkan informasi hasil klasifikasi kinerja asesor BAN PAUD dan PNF Provinsi Sumatera Utara.
- 2) Hasil klasifikasi penilaian kinerja asesor BAN PAUD dan PNF Provinsi Sumatera Utara menggunakan algoritma Naive Bayes memiliki hasil probabilitas tiap hipotesis untuk class tertinggi yaitu class “Baik” berjumlah 0,0047, class “Sangat Baik” berjumlah 0,0002 dan class “Cukup” berjumlah 0,000003.
- 3) Hasil akurasi dengan menggunakan Use Training Set, presentase untuk Correctly Classified Instance yaitu sebesar 97,0799% sementara persentase untuk Incorrectly Classified Instance adalah sebesar 2,9221%. Di mana dari 308 data penilaian kinerja asesor ada sebanyak 299 data asesor yang berhasil diklasifikasikan dengan benar dan

sebanyak 9 data asesor tidak berhasil diklasifikasikan dengan benar..

### UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Allah SWT yang telah memberikan kesehatan dan kelapangan waktu untuk menyelesaikan penelitian ini. Terima kasih juga disampaikan kepada Universitas Potensi Utama khususnya Bapak Edy Victor Haryanto, M.Kom selaku dosen pembimbing yang telah bersedia meluangkan waktunya untuk membimbing penulis dalam menyelesaikan penelitian ini.

### DAFTAR PUSTAKA

- [1] Siddik, M., Hendri, H., Putri, R. N., Desnelita, Y., & Gustientiedina, G. (2020). Klasifikasi Kepuasan Mahasiswa Terhadap Pelayanan Perguruan Tinggi Menggunakan Algoritma Naive Bayes. *INTECOMS: Journal of Information Technology and Computer Science*, 3(2), 162-166.
- [2] Peraturan Menteri Pendidikan dan Kebudayaan Nomor 13 Tahun 2018 tentang Badan Akreditasi Nasional Sekolah/Madrasah Dan Badan Akreditasi Nasional Pendidikan Anak Usia Dini Dan Pendidikan Nonformal. Jakarta.
- [3] Badan Akreditasi Pendidikan Anak Usia Dini dan Pendidikan Nonformal., 2019. Panduan Akreditasi BAN PAUD dan PNF. Jakarta.
- [4] Buulolo, E. (2020). *Data Mining Untuk Perguruan Tinggi*. Deepublish.
- [5] Elisa, E. (2018). Market Basket Analysis Pada Mini Market Ayu Dengan Algoritma Apriori. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2(2), 472-478.
- [6] Novika Sari, V., Yorita Astri, L., & Rasywir, E. (2020). ANALISIS DAN PENERAPAN ALGORITMA NAIVE BAYES UNTUK EVALUASI KINERJA KARYAWAN PADA PT. PELITA WIRA SEJAHTERA. In *Jurnal Ilmiah Mahasiswa Teknik Informatika* (Vol. 2, Issue 1).
- [7] Fitriani, A. S. (2019). Penerapan Data Mining Menggunakan Metode Klasifikasi Naive Bayes untuk Memprediksi Partisipasi Pemilihan Gubernur. *JTAM (Jurnal Teori dan Aplikasi Matematika)*, 3(2), 98-104.
- [8] Kurniawan, Y. I. (2018). Perbandingan Algoritma Naive Bayes dan C. 45 dalam Klasifikasi Data Mining. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, 5(4), 455-464.
- [9] Saputro, I. W., & Sari, B. W. (2020). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Creative Information Technology Journal*, 6(1), 1-11.
- [10] Saputro, I. W., & Sari, B. W. (2020). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Creative Information Technology Journal*, 6(1), 1-11.
- [11] Sadikin, M., Rosnelly, R., & Surya Gunawan, T. (2020). *JURNAL MEDIA INFORMATIKA BUDIDARMA Perbandingan Tingkat Akurasi Klasifikasi Penerimaan Dosen Tetap Menggunakan Metode Naive Bayes Classifier dan C4.5*. 4, 1100–1109. <https://doi.org/10.30865/mib.v4i4.2434>
- [12] Kausar, A., Irawan, A., & Fernando, I. (2023). IMPLEMENTASI ALGORITMA NAIVE BAYES CLASSIFIER UNTUK PENILAIAN KINERJA DOSEN. *PROSISKO: Jurnal Pengembangan Riset dan Observasi Sistem Komputer*, 10(2), 117-127.
- [13] Suardika, I. G. I. (2019). Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naive Bayes: Studi Kasus Fakultas Ekonomi Dan Bisnis Universitas Pendidikan Nasional. *Jurnal Ilmu Komputer Indonesia*, 4(2), 37-44.
- [14] Setiawan, D. A., Halilintar, R., & Wahyuniar, L. S. (2021, August). Penerapan Metode Naive Bayes Untuk Klasifikasi Penentuan Penerima Bantuan PKH. In *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)* (Vol. 5, No. 2, pp. 249-254).
- [15] Rahmadi, M., Kaurie, F., & Susanti, T. (2020). Uji akurasi dataset pasien pasca operasi menggunakan algoritma naive bayes menggunakan weka tools. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 134-139.