

# Algoritma Naive Bayes untuk Menentukan Kelayakan Perpanjangan Kontrak pada PT. Gemilang

<sup>1</sup>Diyah Kiki Widiyaningrum, <sup>2</sup>Riska Septiani, <sup>3</sup>Pamela Larasati, <sup>4</sup>Arief Wibowo  
<sup>1</sup>Magister Ilmu Komputer, Fakultas Teknologi Informasi, Universitas Budi Luhur

Alamat Surat  
Email: [diyah.kikiw@gmail.com](mailto:diyah.kikiw@gmail.com)

## Article History:

Diajukan: 23 Juli 2020; Direvisi: 31 Oktober 2020; Accepted: 11 Nopember 2020

## ABSTRAK

Kerjasama dengan customer merupakan kegiatan rutin yang menjadi salah satu kunci berkembangnya suatu perusahaan. PT Gemilang perusahaan yang bergerak di bidang jasa pelayanan pada sektor telekomunikasi di Indonesia. Untuk dapat terus meningkatkan pelayanan dan mengembangkan kerjasama dengan customer tentunya juga harus didukung dengan cashflow yang baik. Maka dari itu tujuan penulis membuat penelitian ini adalah untuk mengetahui dan mengukur ketepatan waktu pembayaran dari customer atas penyediaan jasa layanan yang disediakan oleh PT. Gemilang dan sebagai salah satu bahan pertimbangan untuk menentukan kelayakan perpanjangan kontrak kerjasama dengan customer dimasa mendatang. Penelitian ini menggunakan data pembayaran customer dari bulan Januari sd Desember 2019 yang kemudian diimplementasikan menggunakan algoritma Naive Bayes. Hasil penelitian yang dilakukan membuktikan bahwa dalam penerapan algoritma naive bayes dapat digunakan dalam menilai kelayakan perpanjangan kontrak pada PT. Gemilang. Tahap pengolahan data yang dilakukan sangat berpengaruh pada nilai hasil accuracy serta evaluasi perbandingan dalam menggunakan aplikasi RapidMiner dan aplikasi Weka. Hasil dari pengolahan data menggunakan aplikasi RapidMiner adalah sebesar 90.62%, sedangkan data yang dihasilkan oleh aplikasi Weka menghasilkan accuracy sebesar 96.8553%. Berdasarkan hasil accuracy yang dihasilkan dapat diambil kesimpulan bahwa kedua aplikasi ini mampu menghasilkan probabilitas ketepatan waktu pembayaran tagihan untuk menentukan kelayakan perpanjangan kontrak kerjasama dengan customer.

**Kata kunci:** Kelayakan Kontrak; Customer; RapidMine; Weka; Naives Bayes

## ABSTRACT

Cooperation with customers is a routine activity that is one of the keys to a company's development. PT Gemilang is a company engaged in service services in the communication sector in Indonesia. To continuously improve service and develop cooperation with customers must also be supported with good cashflow. Therefore, this study aims to know and measure the timeliness of payments from customers for the provision of services provided by PT. Gemilang and as one of the consideration materials to determine the feasibility of contract extension cooperation with customers in the future. The study used customer payment data from January to December 2019, which was then implemented using the Naive Bayes algorithm. The study results proved that the application of a Naive Bayes algorithm could be used to assess the feasibility of contract extension in PT. Resounding. The data processing stage is very influential on the value of accurate results and evaluating comparisons in using RapidMiner applications and Weka applications. The result of processing data using the RapidMiner application is 90.62%, while the data generated by the Weka application generates an accuracy of 96.8553%. Based on the resulting accuracy, results can be drawn to conclude that these two applications can produce a probability of punctuality of bill payment to determine the feasibility of contract extension cooperation with the customer.

**Keywords:** *Unified Modeling Language; Traceability; filing systems; use case*

## 1. PENDAHULUAN

Kontrak merupakan dokumen perjanjian dalam dunia bisnis melakukan kerjasama dengan *customer* merupakan salah satu kegiatan rutin yang menjadi kunci berkembangnya suatu perusahaan. Untuk dapat terus meningkatkan pelayanan dan kerjasama dengan *customer* tentunya juga harus didukung dengan *cashflow* yang baik. Tujuan penulis membuat penelitian ini adalah untuk mengetahui dan mengukur ketepatan waktu pembayaran dari *customer* pada penyediaan jasa layanan oleh PT. Gemilang dan sebagai salah satu bahan pertimbangan untuk menentukan kelayakan perpanjangan kontrak kerjasama dengan customer dimasa mendatang. Dalam penelitian ini, penulis menggunakan data pembayaran dari customer dari bulan januari sd desember 2019 menggunakan metode *Naive Bayes* yang kemudian diimplementasikan menggunakan aplikasi RapidMiner dan aplikasi Weka. PT Gemilang mempunyai salah satu tugas utama untuk mengembangkan beberapa set model dan teknik yang dapat digunakan untuk memprediksi kelayakan perpanjangan kontrak kerjasama dengan customer.

## 2. METODE

Dalam klasifikasinya, penelitian ini menggunakan algoritma *Naive Bayes* yang merupakan metode yang di jabarkan oleh seorang ilmuwan Inggris bernama Thomas Bayes dan juga menjadi salah satu metode untuk klasifikasi dan metode probabilitas serta statistik. Salah satu ciri utama dari metode *Naive Bayes Classifier* ini yaitu asumsi yang sangat kuat dengan independensi dari suatu kondisi atau kejadian. Algoritma *Naive Bayes* dapat memprediksi jumlah peluang di masa yang akan datang. Dalam hal ini yang dimaksud dengan keuntungan penggunaan yaitu dalam metode ini yang diperlukan hanya data pelatihan (*training data*) dengan jumlah kecil yang kemudian digunakan sebagai penentu estimasi parameter sesuai dengan kebutuhan dalam proses klasifikasi yang dilakukan. Karena dianggap sebagai variabel independent, maka dari itu hanya dibutuhkan varians dari suatu variabel pada suatu kelas supaya bisa menjadi penentu dalam proses klasifikasi, bukan jumlah dari keseluruhan matriks kovarians. Teorema Bayes juga sering dikembangkan karena berlaku hukum probabilitas total, sehingga menjadi seperti dibawah ini :

$$P(H|X) = \frac{P(X|H)}{\sum_{i=1}^n P(H_i|X)} \cdot P(H)$$

Keterangan :

- X : Data dengan class yang belum diketahui
- H : Hipotesis data merupakan suatu class spesifik
- P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probabilitas)
- P(H) : Probabilitas hipotesis H (prior probabilitas)
- P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X) : Probabilitas X

Untuk dapat memberikan penjelasan tentang metode tersebut, maka perlu pengklasifikasian dibutuhkan sebanyak petunjuk sebagai penentu kelas apa yang dinilai cocok untuk sampel yang akan dianalisis. Oleh karena hal tersebut, Teorema Bayes diatas disesuaikan seperti berikut ini :

$$P(C|F_1, \dots, F_n) = \frac{P(F_1, \dots, F_n|C)}{P(F_1, \dots, F_n)} \cdot P(C)$$

Dalam hal ini Variabel C memiliki fungsi untuk memberitahukan kelas, sedangkan variabel F1...Fn memiliki fungsi untuk memberitahukan bukti fisis sebagai petunjuk yang diperlukan dalam melakukan proses pengelompokan. Oleh sebab itu, pada rumus tersebut dijelaskan juga bahwa jika peluang pada masuknya sampel suatu karakteristik pada kelas C (Posterior) merupakan peluang munculnya kelas C sebelum sampel masuk, seringnya dinamakan sebagai prior), dikalikan dengan peluang kemunculan dari beberapa bukti fisis dari suatu sampel dalam kelas C (dinamakan likelihood), kemudian dibagi dengan peluang munculnya bukti -bukti fisis sampel secara global (yang juga disebut sebagai *evidence*). Oleh karena itu, rumus di atas dapat pula ditulis secara sederhana sebagai berikut:

$$Posterior = \frac{prior \times likelihood}{evidence}$$

Nilai *Evidence* akan selalu tetap pada masing - masing kelas dalam satu sampel. Nilai dari posterior tersebut yang nantinya akan dibandingkan dengan nilai – nilai pada posterior kelas lainnya fungsinya digunakan sebagai penentu kedalam kelas apa sampel akan diklasifikasikan. Aturan perkalian mengenai rumus Bayes dapat juga dilakukan dengan menjelaskan (C,F1, ... , Fn) dengan rumus sebagai berikut ini :

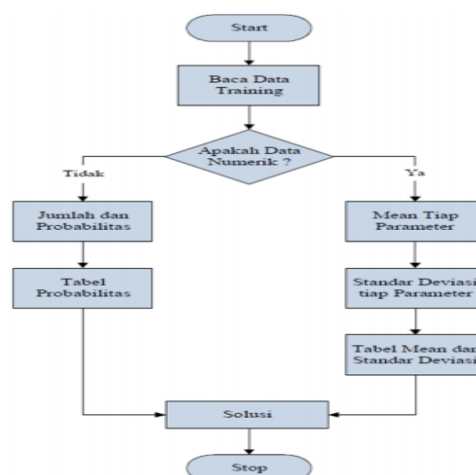
$$\begin{aligned} P(C|F_1, \dots, F_n) &= P(C).P(F_1, \dots, F_n|C) \\ &= P(C).P(F_1|C).P(F_2, \dots, F_n|C, F_1) \\ &= P(C).P(F_1|C).P(F_2|C, F_1).P(F_3, \dots, F_n|C, F_1, F_2) \\ &= P(C).P(F_1|C).P(F_2|C, F_1).P(F_3|C, F_1, F_2) \dots P(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}) \end{aligned}$$

Dari perkalian diatas dapat dilihat bahwa pada hasil penjelasan tersebut mengakibatkan banyaknya faktor lain jika kompleks dan faktor - faktor yang digunakan dalam syarat untuk mempengaruhi nilai pada probabilitas, sehingga mustahil untuk bisa dilakukan analisa satu persatu. Maka perhitungan ini akan berakibat menjadi sulit dilakukan. Dari sinilah kita dapat melihat manfaat yang digunakannya yaitu asumsi independensi yang sangat tinggi (naif), bahwa pada masing - masing petunjuk (F1,F2 , ..., Fn) dapat bebas (*independen*) satu dengan yang lain.

## 2.1 Alur Metode Naïve Bayes

Alur pada metode naïve bayes dapat dilihat dalam gambar berikut ini :

Gambar 1. Alur Metode Naïve Bayes



### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

Penelitian ini menggunakan data pembayaran dari customer pada bulan Januari sd Desember 2019 yang berasal dari perusahaan PT Gemilang. Atribut yang digunakan adalah : Tagihan, Usia Member, Kategori, Jumlah Layanan dan Ketepatan. Selanjutnya dilakukan pengolahan data awal yang di peroleh dari pengumpulan data dari 159 *customer*. Untuk mendapatkan data yang berkualitas maka penulis menggunakan beberapa teknik berikut ini :

1. Data *Cleaning* merupakan sebuah teknik yang berguna untuk menghapus *ouliers* serta untuk menghilangkan data *noise*. Selanjutnya dilakukan pengisian pada nilai - nilai pada data yang tidak lengkap (*missing value*) atau data yang hilang, salah satu keunggulan yang dimiliki oleh algoritma naive bayes tersendiri yaitu mampu menangani data yang tidak lengkap (*missing value*), mampu memperbaiki data – data yang tidak konsisten dan dapat memecahkan redudansi atau berulangnya suatu data yang biasanya disebabkan oleh integrasi data.
2. Data *integration* dan *information*, merupakan suatu teknik yang berfungsi untuk menganalisa data. Nama lain dari proses ini adalah teknik integrasi
3. Data *reduction*, yaitu suatu teknik yang berfungsi untuk mereduksi data set dengan cara mengurangi jumlah pada *attribute* serta *reccord* agar dapat menghasilkan data yang bersifat *informative*.

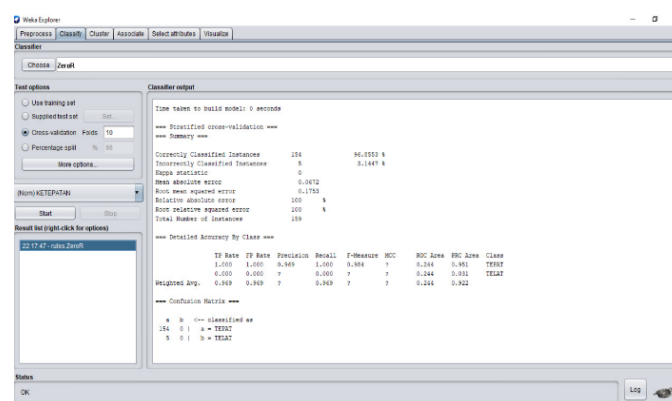
#### PERBANDINGAN HASIL PADA APLIKASI RAPIDMINER DAN WEKA

accuracy: 90.62%

	true TEPAT	true TELAT	class precision
pred. TEPAT	29	0	100.00%
pred. TELAT	3	0	0.00%
class recall	90.62%	0.00%	

Gambar 2. Hasil klasifikasi data menggunakan aplikasi RapidMiner

Dari gambar diatas dapat diambil kesimpulan bahwa metode *naive bayes* yang menggunakan aplikasi RapidMiner menghasilkan ketepatan Tepat sebanyak 29 dan Telat sebanyak 3. Dengan *accuracy* sebesar 90.62%.



Gambar 3. Hasil menggunakan Aplikasi Weka

Dari gambar diatas dapat disimpulkan bahwa hasil klasifikasi dengan metode naive bayes menggunakan aplikasi Weka menghasilkan ketepatan accuracy sebesar 96.8553%. Berdasarkan hasil penelitian yang dilakukan, terbukti bahwa penerapan algoritma naive bayes dapat digunakan untuk menilai kelayakan perpanjangan kontrak pada PT. Gemilang. Tahap pengolahan data sangat

berpengaruh pada nilai hasil accuracy dan evaluasi perbandingan dalam menggunakan aplikasi RapidMiner dan aplikasi Weka. Hasil dari pengolahan data menggunakan aplikasi RapidMiner adalah sebesar 90.62%, sedangkan data yang dihasilkan oleh aplikasi Weka menghasilkan accuracy sebesar 96.8553%. Berdasarkan hasil accuracy yang dihasilkan dapat diambil kesimpulan bahwa kedua aplikasi ini mampu menghasilkan probabilitas ketepatan waktu pembayaran tagihan untuk menentukan kelayakan perpanjangan kontrak kerjasama dengan customer.

#### 4. KESIMPULAN

Dari hasil penelitian diatas dapat dibuktikan bahwa algoritama naive bayes bisa diterapkan untuk menilai kelayakan perpanjangan kontrak pada PT Gemilang. Salah satu hal yang harus diperhatikan adalah tahap pengolahan data, karena proses ini sangat berpengaruh pada nilai hasil accuracy dan evaluasi perbandingan dalam menggunakan aplikasi RapidMiner dan aplikasi Weka. Hasil accuracy dari aplikasi Rapid Miner adalah sebesar 90.62%, sedangkan accuracy yang di hasilkan oleh aplikasi Weka adalah sebesar 96.8553% maka dapat diambil kesimpulan bahwa dari kedua aplikasi ini dapat menghasilkan probabilitas ketepatan waktu customer dalam melakukan pembayaran pada tagihan yang diterima setiap bulannya. Agar bisa dilakukan perbandingan untuk menghasilkan nilai akurasi yang lebih baik dan lebih tinggi nilai maka penulis menyarankan untuk menggunakan 2 (dua) aplikasi, yaitu aplikasi RapidMiner dan Weka.]

#### 5. DAFTAR PUSTAKA

- Jange, B. (2020) "Impacts of Indonesia's 2019 Presidential Election Towards Stock Prices Indonesia Stock Exchange", Jurnal Ilmu Komputer dan Bisnis, 11(1), pp. 2293-2305. Retrieved from <https://ojs.stmikdharmapalariau.ac.id/index.php/jikb/article/view/193> (Accessed: 11October 2020)..
- Nestary, N. (2020). Design of Sales Information Systems in Stock Point Lily Shop is based on PHP MySQL. Jurnal Ilmu Komputer Dan Bisnis, 11(1), 2320-2337. Retrieved from <https://ojs.stmikdharmapalariau.ac.id/index.php/jikb/article/view/195>
- Wongso, F. (2014). Perancangan Sistem Informasi Penerimaan Siswa Baru di SMK Dharma Loka Pekanbaru Berbasis Visual Basic. Jurnal Ilmu Komputer Dan Bisnis, 5(2), 1094-1110. Retrieved from <https://ojs.stmikdharmapalariau.ac.id/index.php/jikb/article/view/49>
- Sahlberg, P. (2012). The most wanted: Teachers and teacher education in Finland. In L. Darling-Hammond & A. Lieberman (Eds.). *Teacher education around the world: changing policies and practices*. London: Routledge, pp. 22-44.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2005). *How People Learn: Brain, Mind, Experience, and School*. <https://www.nap.edu/catalog/9853/how-people-learn-brain-mind-experience-and-school-expanded-edition>.