**PREDIKSI PERPANJANGAN KONTRAK KARYAWAN MENGGUNAKAN TEKNIK DATA MINING NAÏVE BAYES**

## Agung Budiarto1), Bhisma Aldi Pratama2), Avanendra Varrel Sulistya3), Muhammad Fadil4)

Program Studi Sistem Informasi, Universitas Komputer Indonesia

Jl. Dipatiukur No. 112-116, Bandung, 40132, Indonesia

[agung.10519161@mahasiswa.unikom.ac.id](mailto:agung.10519161@mahasiswa.unikom.ac.id)1, [bhisma.10519067@mahasiswa.unikom.ac.id2](mailto:bhisma.10519067@mahasiswa.unikom.ac.id2), [avaneendra.10519070@mahasiswa.unikom.ac.id3](mailto:avaneendra.10519070@mahasiswa.unikom.ac.id3),  [Fadil.10519179@mahasiswa.unikom.ac.id](mailto:agung.10519161@mahasiswa.unikom.ac.id)4.

**Abstrak**

Kontrak kerja adalah suatu perjanjian antara pekerja dan pengusaha secara lisan dan atau tulisan, baik untuk waktu tertentu maupun untuk waktu tidak tertentu yang memuat syarat-syarat kerja, hak dan kewajiban. Penilaian karyawan terkadang masih dinilai kurang objektif, sehingga menimbulkan anggapan bahwa data dapat dimanipulasi dan hasil dari penilaian tersebut menjadi terasa kurang adil bagi beberapa karyawan. Untuk menghadapi permasalahan tersebut, penulis mengambil judul penelitian yaitu “Prediksi Perpanjangan Kontrak Karyawan Menggunakan Teknik Data Mining Naïve Bayes”. Tujuan dari penelitian ini adalah menerapkan metode data mining Naïve Bayes untuk memprediksi perpanjangan kontrak karyawan. Pembuatan perangkat lunak sistem prediksi ini menggunakan bahasa pemrograman Python dengan bantuan software Jupyter Notebook dengan menggunakan dataset yang sudah dibuat sebagai bahan penelitian. Hasil pengujian menghasilkan akurasi 100%. Berdasarkan penelitian yang dilakukan, diperoleh kesimpulan bahwa metode Naïve Bayes dapat digunakan untuk mengklasifikasi data untuk prediksi perpanjangan kontrak dengan akurat.

***Kata Kunci :*** *Data Mining, Prediksi, Naïve Bayes*

## Pendahuluan

Kontrak kerja adalah suatu perjanjian antara pekerja dan pengusaha secara lisan dan atau tulisan, baik untuk waktu tertentu maupun untuk waktu tidak tertentu yang memuat syarat-syarat kerja, hak dan kewajiban. Kontrak kerja dapat berlangsung baik untuk waktu tertentu maupun waktu tidak tertentu. Setiap perusahaan wajib memberikan kontrak kerja dihari pertama anda bekerja. Dalam kontrak kerja biasanya terpapar dengan jelas pekerja memiliki hak mendapat kebijakan perusahaan yang sesuai dengan Undang-Undang Ketenagakerjaan yang berlaku di Indonesia. [1]

PT.X merupakan salah satu perusahaan yang bergerak di bidang penjualan. Sama seperti perusahaan lain pada umunya, PT.X masih menggunakan sistem kontrak kerja karyawan. Setiap setahun periode bekerja pasti ada karyawan yang diperpanjang masa kontraknya ataupun tidak diperpanjang masa kontraknya. PT.X terkadang masih dinilai kurang objektif dalam penilaian karyawan dalam hal perpanjangan kontrak masa kerja karyawannya, sehingga menimbulkan anggapan bahwa data dapat dimanipulasi dan hasil dari penilaian tersebut menjadi terasa kurang adil bagi beberapa karyawan. Selain itu, HRD sering merasa kesulitan dan memerlukan waktu dan proses yang cukup lama. Terlebih lagi, variabel penilaian terkadang masih belum jelas dan karyawan yang banyak dan terus bertambah setiap tahunnya.

Untuk menghadapi permasalahan tersebut, PT.X perlu membuat otomasi perhitungan data dalam bentuk sistem. Maka diperlukannya suatu metode untuk memprediksi apakah karyawan kontrak tersebut diperpanjang atau tidak diperpanjang berdasarkan beberapa variabel penilaian menggunakan teknik data mining klasifikasi.

Klasifikasi adalah proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. [kamber,2006] Klasifikasi Naïve bayes menjadi teknik data mining yang cocok untuk permasalahan ini karena memiliki ketelitian dan kecepatan yang tinggi saat diaplikasikan untuk jumlah data yang besar. Klasifikasi Naïve bayes adalah suatu teknik populer untuk klasifikasi data dalam jumlah yang cukup besar dan probabilitas suatu class.

Selain untuk membantu HRD dalam pengambilan keputusan kontrak kerja karyawan, pentingnya proses perpanjangan kontrak kerja karyawan secara otomatis dan terkomputerisasi juga dapat mempersingkat waktu pekerjaan, sehingga pekerjaan lainnya dapat dilakukan dengan efektif dan efisien. Berdasarkan masalah dari yang sudah dipaparkan, penulis menggunakan teknik klasifikasi menggunakan algoritma Naïve Bayes sebagai metode penyelesaian dari masalah yang ada.

## 1.1 Identifikasi Masalah

Berdasarkan latar belakang masalah yang sudah penulis uraikan, maka identifikasi masalah yang timbpenguul yaitu sebagai berikut:

1. Analisis masih belum objektif sehingga banyak timbul anggapan bahwa data dapat dimanipulasi dan beberapa karyawan merasa tidak adil dengan hal tersebut.
2. Penilaian karyawan belum akurat dan memakan waktu yang cukup lama sehingga pekerjaan tidak efektif dan efisien.

**1.2 Batasan Masalah**

Batasan masalah yang ada pada permasalahan ini yaitu:

1. Metode data mining yang digunakan adalah klasifikasi dengan algoritma Naïve Bayes.
2. Data yang digunakan yaitu dataset sebanyak lebih dari 500 data karyawan PT.X sebagai sample penelitian.
3. Penelitian ini hanya membantu perusahaan dalam menentukan perpanjangan kontrak kerja karyawan yang menghasilkan label berupa “Ya” dan “Tidak”.
   1. **Tujuan Penelitian**

Tujuan penelitian ini adalah merancang dan membuat aplikasi penilaian karyawan sehingga:

1. Menerapkan metode data mining Naïve Bayes untuk memprediksi perpanjangan kontrak karyawan.
2. Mengetahui hasil prediksi perpanjangan kontrak sesuai dengan standar perusahaan.
   1. **Manfaat Penelitian**
3. Bagi Masyarakat

Memberikan pengalaman dalam pengaplikasian teori terhadap kasus permasalahan nyata di perusahaan dan untuk memperoleh pengetahuan baru dalam bidang penelitian dan penulisan karya ilmiah.

1. Bagi Objek Penelitian

Diharapkan dapat membantu perusahaan terutama HRD (*Human Resource Development)* dalam menentukan perpanjangan kontrak kerja karyawan.

## Landasan Teori

**2.1 Data Mining**

*Data mining* merupakan analisis dan pengamatan data yang besar untuk mendapatkan hubungan dan meringkas data agar dapat dimengerti serta dapat digunakan oleh pemilik data (Han and Kamber, 2011). Dalam arti lain *data mining adalah* suatu langkah dalam *Knowledge Discovery In Databases* (KDD). *Knowledge discovery* sebagai suatu proses yang terdiri atas pembersihan data *(data cleaning)*, integrasi data *(data integration)*, pemilihan data *(data selection)*, transformasi data  *(data transformation)*, *data mining*, evaluasi pola *(pattern evaluation)* dan penyajian pengetahuan *(knowledge presentation)*. Jadi *data mining* dapat dikatakan sebagai cara untuk menjelaskan masa lalu dan memprediksi masa depan melalui analisis data. *Data mining* adalah perpaduan dar pembelajaran mesin, statistik, teknologi database dan kecerdasan buatan.

**2.2 Karyawan**

Karyawan kontrak adalah karyawan yang bekerja pada suatu perusahaan dengan kesepakatan dalam bentuk tertulis atau lisan yang dibuat untuk mengikat hubungan antara karyawan dengan pengusaha dalam waktu tertentu maupun waktu tidak tertentu. Kontrak tersebut akan dianggap sah bila kedua belah pihak menyetujui tanpa paksaan. Perjanjian Kerja Waktu Tertentu (PKWT), yaitu perjanjian kerja yang didasari suatu jangka waktu paling lama 2 tahun dan dapat diperpanjang 1 kali untuk jangka waktu maksimal 1 tahun.

**2.3 Klasifikasi**

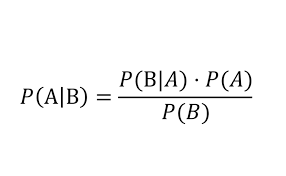
Secara harfiah klasifikasi diartikan sebagai penggolongan atau pengelompokkan. Klasifikasi adalah teknik dalam *data mining* untuk mengelompokkan data berdasarkan keterikatan data terhadap data sampel. Teknik ini dibedakan menjadi beberapa klasifikasi *data mining* yaitu, *tracking pattern/sequencing, classification,assosiation, outlier detection, clustering, regression* dan *forecasting/prediction*. Metode klasifikasi yang sering digunakan yaitu, *Logistic Regression, Naive Bayes, Decision Tree, Random Forest, K-Nearest Neighbour* dan *Artificial Neural Network*.

**2.4 Naïve Bayes**

Bayesian classification adalah suatu proses pengklasifikasian statistik yang dapat digunakan untuk memprediksi kemungkinan atau probabilitas keanggotaan suatu kelas. Naive Bayes adalah metode yang cocok untuk klasifikasi biner dan *multiclass*. Metode ini dikenal sebagai teknik *supervide* klasifikasi objek dimasa depan dengan menetapkan label kela ke instance catatan menggunakan kemungkinan atau probabilitas bersyarat. Kemungkinan atau probabilitas bersyarat adalah ukuran peluang suatu peristiwa yang terjadi berdasarkan peristiwa lain yang telah (dengan asumsi, praduga, pernyataan, atau terbukti) terjadi.

**2.5 Pesamaan Metode Naïve Bayes**

Bentuk umum atau persamaan dari teorema Bayes adalah :



Keterangan:

X : Data dengan kelas yang belum diketahui.

H : Hipotesa data X merupakan suatu class spesifik.

P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probability).

P(H) : Probabilitas hipotesis H (prior probability).

P(X|H): Probabilitas X berdasarkan kondisi pada hipotesis.

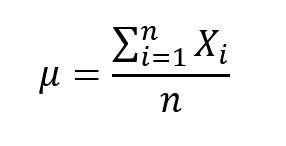
P(X) : Probabilitas X.

**2.6 Alur Metode Naïve Bayes**

Adapun alur metode naive bayes adalah sebagai berikut :

1. Baca *Data Training*
2. Hitung jumlah dan probabilitas, namun apabila data numerik maka :
3. Cari nilai meran dan standar deviasi dari masing-masing parameter yang merupakan data numerik.

Persamaan untuk menghitung nilai rata-rata (mean) adalah sebagai berikut :



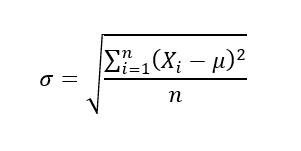
Keterangan :

µ : Rata-rata hitung (mean)

𝑥𝑖 : Nilai sample ke –i

𝑛 : Jumlah sample

Dan persamaan untuk menghitung nilai standar deviasi adalah sebagai berikut:



Keterangan :

σ :standar deviasi

xi : nilai x ke-i

μ : nilai rata-rata hitung (mean)

n : jumlah sampel

1. Cari nilai probabilistic dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.

3. Mendapatkan nilai dalam tabel mean, standar deviasi dan probabilitas.

4. Solusi kemudian dihasilkan.

## Metode

Penelitian ini bertujuan untuk memprediksi perpanjangan kontrak karyawan menggunakan teknik data mining Naive Bayes, aplikasi yang digunakan untuk simulasi adalah Jupyter Notebook yaitu aplikasi pengolah data. Tahapan proses pada riset ini pertama kali adalah penelitian yang terdahulu. Kedua, review literatur. Ketiga, pengumpulan dataset yang akan digunakan. Keempat, analisa terhadap dataset yang digunakan. Kelima, mengimplementasikan dataset dengan model Naive Bayes untuk mengklasifikasi penilaian karyawan.

**3.1 Penelitian Terdahulu**

Penelitian sebelumnya yang telah dilakukan yaitu penelitian oleh Deni Priyanto sebagai mahasiswa program studi Teknik Informatika Universitas Semarang dalam skripsi yang berjudul “Penerapan Data Mining Menggunakan Algoritma Naive Bayes untuk Klasifikasi Perpanjangan Kontrak Kerja Karyawan CV. Jaya Abadi Cleanindo” pada tahun 2020. Hasil dari penelitian tersebut yaitu menghasilkan sistem perpanjangan kontrak kerja karyawan CV. Jaya Abadi CleanIndo menggunakan algoritma Naive Bayes dan pemrograman CodeIgniter dengan akurasi perhitungan sebesar 90%.

**3.2 Review Literatur**

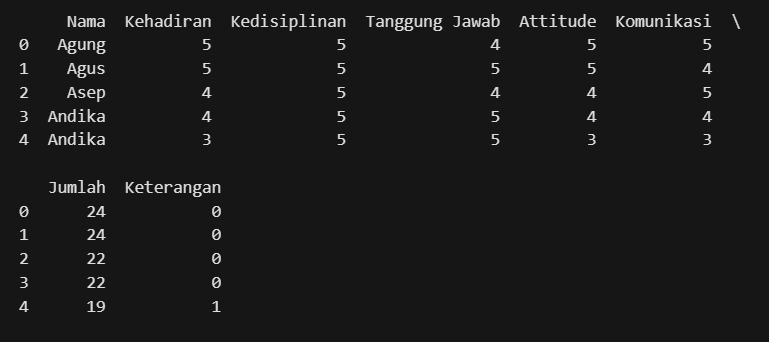
Pada bagian ini dilakukan review literatur dari bidang ilmu yang berkaitan pada penelitian ini diantaranya adalah :

1. Data Mining
2. Naive Bayes

Literatur didapatkan dari jurnal, buku, maupun penelitian sebelumnya dan artikel yang relevan dengan penelitian ini juga menggunakan metode yang sama.

**3.3. Pengumpulan Dataset**

Dataset yang digunakan dalam penelitian ini ialah berdasarkan data dari dataset yang kami buat dari PT. X. Format dari dataset ini CSV yang terdiri dari 641 baris data dan 6 kolom data. Setiap baris mewakili atribut Kehadiran, Kedisiplinan, Tanggung Jawab, Attitude, Komunikasi, dan Hasil.



Gambar 1Dataset

**3.4 Analisis Dataset**

Pada riset ini menggunakan metodologi pengembangan data mining CRISP- DM( Cross Standart Industries for Data Mining) ialah merupakan standar yang diperuntukan guna proses analisis sesuatu industry selaku strategi pemecahan permasalahan dari bisnis., yang mempunyai siklus hidup terdiri dari 6 sesi.

**3.5 Implementasi Naïve Bayes**

Naive Bayes adalah teknik pengklasifikasian prediksi yang berbasis probabilitas sederhana bersumber pada pelaksanaan teorema Bayes dengan anggapan independensi yang kuat dengan menghitung frekuensi dan campuran dari nilai dataset yang ada. Naive Bayes menggunakan fitur independen yang kuat dengan kata lain, bahwasannya data tidak berkaitan pada data lain dalam kasus yang sama ataupun atribut lain.

Contoh dari rumus persamaan Teorema Bayes adalah :

**P(H│X)=(P(X│H).P(H))/(P(X))**

Di mana : 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

## 4. Hasil

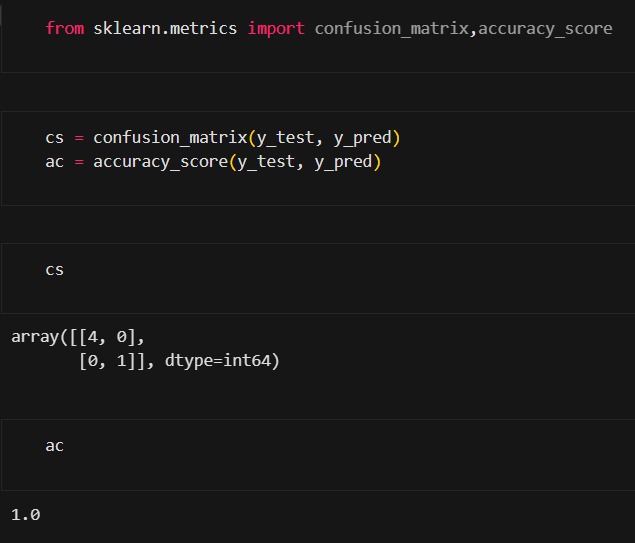
Atribut penilaian yang digunakan untuk memprediksi perpanjangan kontrak karyawan PT. X.

Table 1Atribut Penilaian

|  |  |  |
| --- | --- | --- |
| No | Atribut | Nilai |
| 1. | Absensi | 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 2. | Kinerja | 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 3. | Pemahaman | 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 4. | Attitude | 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 5. | Komunikasi | 1 |
|  | 2 |
|  | 3 |
|  | 4 |
|  | 5 |

Hasil perpanjangan kontrak karyawan didapat dari total nilai atribut penilaian. Jika total atribut nilai karyawan kecil dari 20 maka kontrak karyawan tidak diperpanjang. Sedangkan, jika total atribut nilai karyawan lebih dari 20 makan kontrak karyawan diperpanjang.

Pemodelan algoritma Naïve Bayes yang dihasilkan berikut gambarnya dapat dilihat dibawah ini.



Gambar 2Hasil

Berdasarkan hasil pemodelan Naïve Bayes diatas machine learning ini dapat memberikan hasil yang benar dan akurat sebesar 100%, didapatkan hasil bahwa karyawan tersebut kontrak nya diperpanjang karena memenuhi syarat total nilai atribut untuk perpanjangan kontrak karyawan di PT. X.

## Kesimpulan

Kesimpulan yang didapat yaitu dengan menggunakan model Naïve Bayes bisa membantu penilaian layak atau tidaknya karyawan dalam mendapat perpanjangan kontrak di PT. X. Tingkat akurasi untuk penilaian kelayakan karyawan mendapat perpanjangan kontrak di PT. X sebesar 100% yang dapat membantu pimpinan dalam memberi keputusan.

## Saran

Beberapa saran yang mungkin dapat dipertimbangkan untuk pengembangan selanjutnya yaitu untuk kedepannya semoga dapat menambahkan pemodelan lainnya untuk mendapatkan akurasi yang lebih baik.

## Daftar Pustaka

Fadlan, C., Ningsih, S., & Windarto, A. P. (2018). Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra. *JUTIM (Jurnal Teknik Informatika Musirawas)*, *3*(1), 1-8.

WK, W. N., & Da Cintia, F. (2018).

IMPLEMENTASI DATA MINING UNTUK PREDIKSI STATUS KONTRAK KERJA KARYAWAN MENGGUNAKAN ALGORITMA NAIVE BAYES STUDI KASUS KOSPIN JASA. *Jurnal Komputer Bisnis*, *11*(2), 23-32.

Feldman,R., and J. Sanger. 2007. The Text Mining Handbook :Advanced Approaches in Analysing Unstructured Data. New York : Cambridge University Press

El Kourdi et al. 2004. Automatic Arabic Document Categorization

Based on the Naïve Bayes Algorithm,AlakhawaynUniversity,Marocco,http://www.aclweb.org/anthology/W04-1610.pdf, 2 Mei 2015

WK, W. N., & Da Cintia, F. (2018). IMPLEMENTASI DATA MINING UNTUK PREDIKSI STATUS KONTRAK KERJA KARYAWAN MENGGUNAKAN ALGORITMA NAIVE BAYES STUDI KASUS KOSPIN JASA. *Jurnal Komputer Bisnis*, *11*(2), 23-32.

WK, W. N., & Da Cintia, F. (2018). IMPLEMENTASI DATA MINING UNTUK PREDIKSI STATUS KONTRAK KERJA KARYAWAN MENGGUNAKAN ALGORITMA NAIVE BAYES STUDI KASUS KOSPIN JASA. *Jurnal Komputer Bisnis*, *11*(2), 23-32.

Winarti, T., Priyanto, D., Vydia, V., & Indriyawati, H. (2020). Penerapan Data Mining Menggunakan Algoritma Naive Bayes Untuk Klasifikasi Perpanjangan Kontrak Kerja Karyawan. In *Seminar Nasional Hasil Penelitian dan Pengabdian Kepada Masyarakat* (pp. 288-301).

Susanto, H., & Sudiyatno, S. (2014). Data mining untuk memprediksi prestasi siswa berdasarkan sosial ekonomi, motivasi, kedisiplinan dan prestasi masa lalu. *Jurnal pendidikan vokasi*, *4*(2).

Danny, M., & Yusup, D. (n.d.). ANALISIS DATA MINING MENGGUNAKAN ALGORITMA NAÏVE BAYES DALAM MEMPREDIKSI . SIGMA Jurnal Teknologi Pelita Bangsa.

Abdul Muiz Khalimi, S., 2022. Menghitung Naive Bayes.