

IMPLEMENTASI DATA MINING UNTUK PREDIKSI STATUS KONTRAK KERJA KARYAWAN MENGGUNAKAN ALGORITMA NAIVE BAYES STUDI KASUS KOSPIN JASA

Wahyu Nurjaya WK, S.T., M.Kom¹, Frika Da Cintia²

^{1,2}Program Studi Teknik Informatika STMIK LPKIA

Jln. Soekarno Hatta No. 456 Bandung 40266, Telp. +62 22 75642823, Fax. +62 22 7564282

¹wahyunwk@lpkia.ac.id, ²frika6314062@fellow.lpkia.ac.id

Abstrak

Kospin Jasa (Koperasi Simpan Pinjam Jasa) merupakan perusahaan terkemuka berskala nasional yang bergerak di bidang Simpan Pinjam dan Keuangan yang berdiri sejak tahun 1973. Pada saat ini, Kospin Jasa memiliki 286 kantor cabang dan pelayanan dengan jumlah karyawan kontrak sekitar 1.592 orang. Status kerja karyawan kontrak di sebuah perusahaan merupakan hal yang penting dan perlu diperhatikan demi tercapainya tujuan perusahaan. Setiap tahun, perusahaan akan memperpanjang ataupun memutus kontrak kerja karyawan karena siklus tersebut sudah menjadi peraturan dalam sebuah perusahaan. Dengan banyaknya kantor cabang dan karyawan yang dimiliki oleh kospin jasa, maka akan banyak pula karyawan yang akan mengajukan perpanjangan kontrak di tiap tahunnya, sehingga akan membutuhkan waktu yang banyak dalam melakukan penentuan status perpanjangan kontrak kerja karyawan. Berdasarkan hasil pengamatan yang dilakukan di Koperasi Simpan Pinjam JASA Bandung, disimpulkan bahwa proses menentukan status karyawan kontrak akan diperpanjang masa kerjanya belum dapat ditangani secara baik. Kospin Jasa membutuhkan suatu aplikasi yang dapat menangani proses penentuan status perpanjangan kontrak kerja karyawan kontrak. Dengan sistem yang terkomputerisasi dapat mempermudah dan mengurangi waktu dalam proses penentuan masa kontrak kerja karyawan. Pembuatan perangkat lunak sistem prediksi status kontrak kerja karyawan ini akan menggunakan suatu prototipe perangkat lunak yang disebut Perangkat Lunak Prediksi Status Kontrak Kerja Karyawan Kospin Jasa, yang dibangun dengan bahasa pemrograman java menggunakan Netbeans IDE 7.3.1 dan MySQL. Perangkat Lunak ini juga menerapkan algoritma yang disebut *Naive Bayes*.

Kata kunci : *naive bayes*, karyawan kontrak, kospin jasa, status kontrak

1.1. Latar Belakang Masalah

Kospin Jasa (Koperasi Simpan Pinjam Jasa) merupakan perusahaan terkemuka berskala nasional yang bergerak di bidang Simpan Pinjam dan Keuangan yang berdiri sejak tahun 1973. Kospin Jasa memiliki banyak cabang yang tersebar di wilayah Indonesia. Kospin jasa memiliki jumlah nasabah yang sangat banyak dan terus bertambah setiap tahunnya. Pada saat ini, Kospin Jasa memiliki 286 kantor cabang dan pelayanan dengan jumlah karyawan kontrak sekitar 1.592 orang

Karyawan adalah Sumber Daya Manusia (SDM) yang merupakan aset perusahaan yang sangat berharga dan harus dikelola dengan baik agar dapat memberikan kontribusi yang optimal dalam kemajuan suatu perusahaan atau instansi (Fachrizal & Ginting, 2013). Karyawan terdiri dari 2 jenis yaitu karyawan tetap dan karyawan tidak tetap atau yang lebih dikenal dengan karyawan kontrak. Karyawan kontrak adalah karyawan yang bekerja pada suatu instansi dengan kerja waktu tertentu yang didasari atas suatu perjanjian atau kontrak dapat juga disebut dengan Perjanjian Kerja Waktu Tertentu (PKWT),

yaitu perjanjian kerja yang didasarkan suatu jangka waktu yang diadakan untuk paling lama 2 tahun dan hanya dapat diperpanjang 1 kali untuk jangka waktu maksimal 1 tahun (Undang-Undang RI Ketenagakerjaan 2003 dalam pasal 59 ayat 1).

Status kerja karyawan kontrak di sebuah perusahaan merupakan hal yang penting dan perlu diperhatikan demi tercapainya tujuan perusahaan. Setiap tahun, perusahaan akan memperpanjang ataupun memutus kontrak kerja karyawan kontrak karena siklus tersebut sudah menjadi peraturan dalam sebuah perusahaan. Dengan banyaknya kantor cabang dan karyawan yang dimiliki oleh kospin jasa, maka akan banyak pula karyawan yang akan mengajukan perpanjangan kontrak di tiap tahunnya, sehingga akan membutuhkan waktu yang banyak dalam melakukan penentuan status perpanjangan kontrak kerja karyawan. Dalam menentukan apakah seorang karyawan layak untuk diperpanjang atau tidak diperpanjang kontrak kerjanya seringkali menjadi hal yang sulit, disamping banyaknya kriteria penilaian yang

diperlukan untuk memutuskan, faktor lain seperti sifat subjektif yang dilakukan oleh pimpinan cabang dalam memberikan status rekomendasi perpanjangan kontrak karyawan sering kali terjadi dan ini pun menjadi hal yang perlu diperhatikan. Penilaian kinerja karyawan kontrak selalu dilakukan disetiap semester, bagian Sumber Daya Manusia (SDM) hanya menerima formulir pengajuan perpanjangan kontrak beserta hasil penilaian kinerja karyawan dan rekomendasi dari masing masing wakil pimpinan cabang.

Untuk menghadapi permasalahan yang ada, maka diperlukan suatu metode untuk memprediksi apakah karyawan kontrak tersebut akan diperpanjang atau diputus status kontrak kerjanya. Untuk mengatasi hal tersebut, maka digunakanlah teknik prediksi atau peramalan. Teknik data mining yang digunakan dalam penelitian ini adalah algoritma Naive Bayes, teknik ini diharapkan dapat menghasilkan suatu pengetahuan yang dapat digunakan dalam pengambilan keputusan terhadap status kerja karyawan kontrak.

1.2. Identifikasi Permasalahan

Adapun permasalahan atau persoalan yang terjadi berdasarkan latar belakang masalah yang telah dipaparkan, masalah yang dapat diidentifikasi adalah proses penentuan status perpanjangan kontrak karyawan yang membutuhkan waktu yang lama dan belum sesuai dengan prosedur yang ada.

Untuk menghindari pembahasan masalah yang terlalu meluas maka pada penulisan laporan ini dibatasi dengan ruang lingkup sebagai berikut:

1. Penelitian ini hanya berfokus pada status kontrak kerja untuk karyawan kontrak di Kospin Jasa
2. Penelitian ini hanya berfokus pada proses bisnis untuk menetapkan status kontrak kerja karyawan di Kospin Jasa

1.3. Ruang Lingkup Permasalahan

Berdasarkan identifikasi permasalahan yang telah diuraikan sebelumnya, tujuan yang ingin dicapai adalah untuk memudahkan proses penentuan status karyawan kontrak akan diperpanjang atau tidak masa kerjanya.

2.1. Landasan Teori

2.1.1. Data Mining

Saat ini banyak sekali perusahaan yang memanfaatkan teknologi informasi untuk mendukung kegiatan proses bisnis mereka, dan tidak menutup kemungkinan data yang dihasilkan sangat banyak, perkembangan internet juga memiliki andil yang cukup besar dalam akumulasi data perusahaan. Data juga merupakan aset terpenting yang dimiliki perusahaan untuk mencapai tujuan perusahaan di masa yang akan datang. Tetapi pertumbuhan data yang sangat pesat dapat membuat kondisi yang

sering disebut rich of data but poor of information dimana data yang terkumpul tidak lagi dapat digunakan untuk aplikasi yang masih digunakan perusahaan, namun dengan data mining data masa lalu yang sudah tidak digunakan lagi dapat menghasilkan informasi baru yang bermanfaat untuk masa depan. Berikut adalah beberapa definisi data mining dari beberapa penulis, Data mining menurut Aries Saifudin dalam jurnal Metode Data Mining Untuk Seleksi Calon Mahasiswa Pada Penerimaan Mahasiswa Baru Di Universitas Pamulang “Data mining adalah proses untuk menemukan pola yang berguna dan kecenderungan di dalam kumpulan data yang besar.” (Saifudin, 2018).

Selain itu, menurut Alfa Saleh dalam jurnal Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga “Data Mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data berukuran besar.” (Saleh, 2015).

Dan dalam jurnal Data Mining Analisa Pola Pembelian Produk Dengan Menggunakan Metode Algoritma Apriori menurut Heroe Santoso, I Putu Hariyadi, dan Prayitno

“Data mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar.” (Santoso, Hariyadi, & Prayitno, 2016).

Merujuk beberapa definisi data mining dari beberapa penulis diatas, penulis mendefinisikan bahwa dalam penelitian ini Data Mining adalah suatu proses menggali atau mengumpulkan kembali data penilaian kinerja karyawan atau data pengajuan perpanjangan kontrak di masa lalu dalam jumlah yang besar, lalu menemukan gambaran aturan yang diterapkan di data tersebut, untuk mendapatkan informasi baru yang berguna di masa depan.

Untuk melakukan proses data mining harus melalui beberapa tahapan, tahapan tersebut dimulai dari membersihkan data, mengintegrasikan data, menyeleksi data, mentransformasikan data, kemudian masuk ke proses mining data yang akan menghasilkan pola yang ada di dalam data, dan terakhir akan menghasilkan pengetahuan baru yang bermanfaat. Untuk lebih jelas berikut adalah penjelasan mengenai tahap-tahap Data mining :

- a. Pembersihan Data (data cleaning) merupakan tahap untuk memilah dan menghilangkan data yang tidak konsisten, data yang tidak relevan, data null atau data yang kosong dan menyamakan dengan format database yang baru
- b. Integrasi Data (data integration) merupakan tahap untuk menyatukan data dari berbagai sumber database seperti data yang disimpan di file excel, file access, dll, ke dalam database baru

- c. Seleksi Data (Data Selection) merupakan tahap untuk memilah atau menyeleksi data yang akan diambil dari database untuk dianalisis
- d. Data Transformation merupakan tahap untuk mengubah atau menggabungkan data yang diambil dari database untuk menyesuaikan format yang akan diproses dalam data mining
- e. Proses mining merupakan tahap utama ketika metode digunakan untuk menghasilkan pengetahuan baru yang berharga dan tersembunyi dari data.
- f. Evaluasi pola (pattern evaluation) merupakan tahap untuk mengidentifikasi pola-pola menarik kedalam knowledge based yang ditemukan bermanfaat.
- g. Presentasi pengetahuan merupakan tahap visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining.

Dalam data mining terdapat beberapa algoritma untuk menemukan pengetahuan baru, ada lima fungsi utama dari data mining yaitu :

1. Estimasi (Estimation) yaitu untuk menerka nilai yang belum diketahui, misalnya menerka waktu pengiriman pizza. Metode yang biasa digunakan dalam menentukan estimasi yaitu Linear Regression, Neural Network, Support Vector Machine, dll.
2. Prediksi/Peramalan (Prediction/Forecasting) yaitu memperkirakan nilai untuk masa yang akan datang, misalnya memperkirakan harga saham. Metode yang biasa digunakan dalam menentukan prediksi atau peramalan yaitu Neural Network, Decision Tree, K-Nearest Neighbor, dll
3. Klasifikasi (Classification) yaitu proses menemukan model atau fungsi yang menjelaskan pengelompokan sesuatu berdasarkan kelas-kelas tertentu, yang bertujuan untuk memperkirakan kelas dari suatu objek yang labelnya belum diketahui. Metode yang biasa digunakan dalam menentukan klasifikasi yaitu Naive Bayes, K-Nearest Neighbor, C4.5, ID3, CART, Linear Discriminant Analysis, Logistic Regression, dll.
4. Klastering (Clustering) yaitu pengelompokan identifikasi data yang mempunyai karakteristik tertentu. Metode yang biasa digunakan untuk klastering yaitu K-Means, K-Medoids, Self-Organizing Map (SOM), Fuzzy C-Means, dll.
5. Asosiasi (Association) yaitu mengidentifikasi item-item produk yang mungkin akan dibeli oleh konsumen bersamaan dengan produk yang lainnya. Metode yang biasa digunakan untuk asosiasi yaitu FP-Growth, A Priori, Coefficient of Correlation, Chi Square, dll.

2.1.2. Algoritma Naïve Bayes

Naïve Bayes merupakan metode klasifikasi yang sering digunakan untuk memperkirakan suatu nilai yang labelnya belum diketahui. Penggunaan metode Naïve Bayes membutuhkan jumlah data pelatihan atau yang sering disebut data training yang kecil yang digunakan untuk menentukan estimasi parameter yang dibutuhkan dalam proses klasifikasi. Metode Naïve Bayes juga merupakan salah satu metode yang menggunakan perhitungan peluang dari satu kelas dari masing-masing kelompok atribut yang ada, dan menentukan kelas mana yang paling optimal. Berikut beberapa definisi algoritma Naïve Bayes menurut beberapa penulis. Menurut Febri Liantoni dan Hendro Nugroho dalam jurnal *Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan Knearest Neighbor* definisi *Naïve Bayes* adalah sebagai berikut :
 “*Naïve Bayes* adalah metode Bayesian Learning yang paling cepat dan sederhana. Hal ini berasal dari teorema Bayes dan hipotesis kebebasan, menghasilkan klasifier statistik berdasarkan peluang.” (Liantoni & Nugroho, 2015)

Sedangkan menurut informatikalogi.com Naïve Bayes memiliki definisi sebagai berikut :

“Algoritma Naive Bayes merupakan sebuah metoda klasifikasi menggunakan metode probabilitas dan statistik yg dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dr Naïve Bayes Classifier ini adalah asumsi yg sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian.” (informatikalogi, 2017)

Algoritma Naïve Bayes memiliki tahapan dari proses yang harus dilakukan yaitu :

1. Menghitung jumlah kelas / label.
 2. Menghitung Jumlah Kasus Per Kelas
 3. Kalikan Semua Variable Kelas
 4. Bandingkan Hasil Per Kelas
- Berikut adalah rumus persamaan dari teorema Bayes :

$$P(H | X) = (P(X | H) \cdot P(H)) / (P(X))$$

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

Dalam penelitian ini poin poin yang akan digunakan sebagai kriteria atau indikator dalam implementasi metode Naïve Bayes adalah poin poin penilaian kinerja karyawan yang ada di Kospin Jasa, poin poin tersebut antara lain :

1. Pelayanan prima adalah usaha yang dilakukan oleh seluruh karyawan Kospin Jasa untuk memberikan pelayanan yang terbaik kepada nasabah, calon anggota, maupun anggota yang ada di Kospin Jasa. Penilaian pelayanan prima ini dilakukan oleh pejabat penilai yang ada di kantor pusat maupun seluruh kantor cabang Kospin Jasa, dengan cara melihat pelayanan yang diberikan seluruh karyawan Kospin Jasa kepada nasabah setiap harinya.
2. Key performance indicators (KPI) adalah indikator yang merupakan kunci ukuran keberhasilan kinerja suatu bagian yang ada di perusahaan atau di perusahaan tersebut untuk memenuhi tujuan yang telah ditetapkan. Kospin Jasa memiliki KPI yang berbeda di setiap bagian yang ada, misalnya untuk bagian Back Office penilaian KPI memiliki tujuh poin yaitu pertumbuhan produk simpanan, performa karyawan dalam pelayanan, OHC, jumlah transaksi penjournalan, nilai kepatuhan, penguasaan bagian, dan hasil tes pengetahuan karyawan. Penilaian KPI ini dilakukan oleh sistem setiap 6 bulan.
3. Performance Appraisal (PA) adalah penilaian kinerja karyawan berdasarkan prestasi kerja, pada Kospin Jasa penilaian performance appraisal ini dinilai berdasarkan kedisiplinan, kerapian, absensi, dan pengetahuan karyawan akan produk produk yang ada di Kospin Jasa, produk produk yang dimaksud adalah layanan yang disediakan oleh Kospin Jasa.
4. Rekomendasi adalah pernyataan yang menyarankan seorang karyawan layak untuk diperpanjang masa kontrak kerjanya dan diberikan secara langsung oleh pimpinan cabang/kantor.

2.1.3. Kontrak Kerja

Kontrak kerja atau perjanjian kerja merupakan hal yang penting bagi karyawan kontrak dan harus selalu diperhatikan agar tercapainya tujuan perusahaan. Kontrak kerja atau perjanjian kerja juga diatur dalam undang-undang. Berikut beberapa definisi perjanjian kerja atau kontrak kerja menurut undang-undang

“Menurut Undang-Undang No 13 Tahun 2003 tentang Ketenagakerjaan bahwa Kontrak Kerja/Perjanjian Kerja adalah perjanjian antara pekerja/buruh dengan pengusaha atau pemberi kerja yang memuat syarat syarat kerja, hak, dan kewajiban para pihak.” (Shanto, 2017)

Dan menurut Kitab Undang-Undang Hukum Perdata Bab VIIa tentang Perjanjian Kerja Bagian 1 Pasal 1601a perjanjian kerja adalah sebagai berikut :

“Perjanjian kerja ialah suatu persetujuan bahwa pihak kesatu, yaitu buruh, mengikatkan diri untuk menyerahkan tenaganya kepada pihak lain, yaitu majikan, dengan upah selama waktu yang tertentu.”

3.1 Analisis Fungsional

3.1. Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional adalah pernyataan fungsi-fungsi atau proses apa saja yang harus ada di perangkat lunak yang akan dibuat. Dalam penelitian ini, kebutuhan fungsional yang menggambarkan fungsi atau proses yang harus ada di perangkat lunak antara lain :

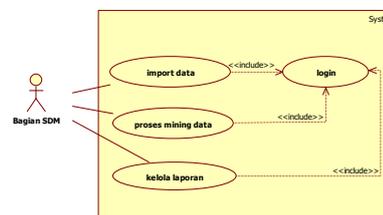
1. Import data dari file csv
2. Proses *mining* data
3. Tampil hasil prediksi berdasarkan cabang tertentu, NIK (Nomor Induk Karyawan) tertentu, dan hasil prediksi keseluruhan dengan bentuk laporan.

Selain analisis kebutuhan fungsional, perangkat lunak yang akan dibuat juga memiliki analisis kebutuhan non fungsional, analisis kebutuhan non fungsional adalah pernyataan batasan fungsi yang akan ada di perangkat lunak, seperti batasan waktu, batasan pengembangan proses, standarisasi perangkat lunak, dll. Dalam penelitian ini, kebutuhan non fungsional yang ada di perangkat lunak antara lain :

1. Bahasa pemrograman menggunakan java
2. Metode yang diterapkan yaitu metode naive bayes
3. Database yang digunakan di perangkat lunak menggunakan MySQL
4. Laporan hasil prediksi yang dihasilkan perangkat lunak akan ditampilkan menggunakan iReport

3.2. Usecase Diagram

Dalam sub bab ini, use case diagram akan mendeskripsikan fungsi-fungsi maupun aktor dari perangkat lunak yang akan dibuat. Berikut adalah rancangan diagram use case dari perangkat lunak yang akan dibuat :



Gambar 1 Use Case Diagram

Tabel 1 Uraian Usecase Scenario Import Data

No	USC-02
Nama Use Case	Import data
Aktor	Bagian SDM
Kondisi Awal	Form import data ditampilkan

Kondisi Akhir	Tabel menampilkan data testing yang telah di import
Aksi Aktor	Reaksi Sistem
Skenario Normal	
1. Klik button Choose File	
	2. Menampilkan file explorer
3. Pilih file (.csv) yang akan diimport	
	4. Menampilkan lokasi file (.csv) yang dipilih
5. klik button Import Data	
	6. memproses import data ke database
	7. menampilkan data dari file yang diimport ke tabel yang disediakan
	8. membersihkan data yang masih <i>null</i> dan data ganda
	9. mengubah format data penilaian kinerja menjadi nilai huruf
	10. menampilkan data testing dari file yang diimport ke tabel yang disediakan

Tabel 2 Uraian Usecase Scenario Proses Mining Data

No	USC-03
Nama Use Case	Proses mining data
Aktor	Bagian SDM
Kondisi Awal	Form import data ditampilkan
Kondisi Akhir	Menampilkan hasil prediksi dari proses data testing
Aksi Aktor	Reaksi Sistem
Skenario Normal	
1. Klik button Proses Data Testing	
	2. Cek data testing
	3. Menghitung jumlah data keseluruhan
	4. Menghitung jumlah data per kelas
	5. Menghitung jumlah data untuk kriteria

	elayanan prima per kelas
	6. Menghitung jumlah data untuk kriteria key performance indicators per kelas
	7. Menghitung jumlah data untuk kriteria performance appraisal per kelas
	8. Menghitung jumlah data untuk kriteria rekomendasi per kelas
	9. Mengalikan semua jumlah data per kriteria untuk setiap kelas
	10. Hasil probabilitas akhir tiap kelas
	11. Membandingkan kelas diperpanjang dan tidak diperpanjang
	12. Cek apakah data diperpanjang atau tidak diperpanjang
	13. Menampilkan hasil prediksi dari data testing

3.3. Tahapan Knowledge Discovery in Database

Pada bab sebelumnya telah dijelaskan bahwa sebelum melakukan proses data mining ada beberapa tahap yang harus dilakukan yaitu tahapan Knowledge Discovery in Database (KDD), dimana tahapan tersebut dimulai dari tahap *selection*, *preprocessing*, *transformation*, *data mining*, dan *interpretation & evaluation*. Untuk lebih jelas berikut adalah penjelasan mengenai tahap-tahap KDD yang dilakukan pada penelitian ini :

A. Selection

Pada tahapan *selection* ini, data awal yang diimport akan langsung masuk ke dalam tahapan pemilihan data training dan data testing. Pada tahap ini lah data awal tersebut akan dipisahkan berdasarkan tahun akhir kontrak. Data testing adalah data pengajuan perpanjangan kontrak kerja karyawan kontrak yang masa kerjanya akan habis dalam tahun ini, sedangkan data training adalah data pengajuan perpanjangan kontrak kerja karyawan yang masa akhir kontraknya dibawah tahun ini. Beberapa kriteria yang akan digunakan sebagai atribut untuk memprediksi status perpanjangan kontrak adalah penilaian kinerja yaitu nilai Pelayanan Prima, Key Performance Indicators, Performance Appraisal, dan Rekomendasi.

B. Preprocessing

Tahap *preprocessing* adalah tahapam yang dilakukan setelah tahap *selection*, pada tahapan ini data yang telah dipisahkan menjadi data *testing* dan

data *training* akan dibersihkan dari data yang tidak sesuai dengan kriteria data yang akan digunakan. Pada penelitian ini tahapan *preprocessing* adalah menghilangkan data yang masih kosong pada field tertentu, dan menghilangkan duplikasi data pada pengajuan perpanjangan kontrak karyawan.

C. Transformation

Tahapan ini dilakukan setelah tahap sebelumnya selesai dilakukan, tahap *transforming* pada penelitian ini adalah mengubah beberapa data dan menyamakan data menjadi format yang sesuai dengan data yang digunakan untuk proses *data mining*. Adapapun data yang harus di *transforming* antara lain:

1. Menyusun data yang didapat menjadi file baru dalam format .csv
2. Membuat tabel sesuai dengan kriteria atau atribut yang digunakan dalam proses prediksi status perpanjangan kontrak yaitu nilai pelayanan prima, *key performance indicators*, *performance appraisal*, dan rekomendasi.
3. Mengubah nilai angka penilaian pelayanan prima, *key performance indicators*, dan *performance appraisal* menjadi nilai grade. Adapun range dari nilai yang akan diubah menjadi grade seperti berikut ini :

Tabel 3 Range skala penilaian

Nilai Angka	Grade	Contoh
> 90 - 100	A	94,20 → A
> 80 - 90	B	87,50 → B
> 60 - 80	C	70, 55 → C
> 40 - 60	D	52, 10 → D
> 0 - 40	E	33,40 → E

D. Data Mining

Jika tahapan *transforming* selesai dilakukan, maka tahapan selanjutnya adalah tahap *data mining*. Tahap *data mining* merupakan tahapan pertama yang dilakukan saat metode diterapkan untuk menemukan pola yang tepat dari data training, pada penelitian ini penulis menggunakan algoritma *Naïve Bayes*. Dalam penerapan *naive bayes* ada beberapa tahap yang harus dilakukan, yaitu:

1. Menghitung jumlah kelas / label.
2. Menghitung Jumlah Kasus Per Kelas
3. Kalikan Semua Variable Kelas
4. Bandingkan Hasil Per Kelas

Untuk contoh perhitungan dari algoritma *naïve bayes* akan dijelaskan di sub bab III.8 Analisis Perhitungan *Naïve Bayes*.

E. Interpretation & Evaluation

3.4. Analisis Perhitungan Naïve Bayes

Analisis perhitungan pada sub bab ini dibuat mulai dari perhitungan prediksi menggunakan tahap-tahap

yang ada pada algoritma *naïve bayes* sampai dengan menghitung akurasi dari hasil prediksi. Data testing yang akan dihitung menggunakan tahapan perhitungan *naïve bayes* dapat dilihat di tabel 3.3 dibawah ini, dalam tabel data testing tersebut ada 10 data penilaian kinerja karyawan, adapun 4 tahapan perhitungan *naïve bayes* dimulai dari menghitung jumlah kelas per label, menghitung jumlah kasus per kelas, mengalikan semua variable kelas, dan yang terakhir membandingkan hasil per kelas, perhitungan lebih detailnya adalah seperti berikut :

Tabel 4 Data Testing

No	Nama	Abad	Jenis Kelamin	Usia	Keputusan	Keputusan	Keputusan	Keputusan	Keputusan	Keputusan
1	Agung	Male	Male	30	Ya	Ya	Ya	Ya	Ya	Ya
2	Agung	Male	Male	30	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak
3	Agung	Male	Male	30	Ya	Ya	Ya	Ya	Ya	Ya
4	Agung	Male	Male	30	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak
5	Agung	Male	Male	30	Ya	Ya	Ya	Ya	Ya	Ya
6	Agung	Male	Male	30	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak
7	Agung	Male	Male	30	Ya	Ya	Ya	Ya	Ya	Ya
8	Agung	Male	Male	30	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak
9	Agung	Male	Male	30	Ya	Ya	Ya	Ya	Ya	Ya
10	Agung	Male	Male	30	Tidak	Tidak	Tidak	Tidak	Tidak	Tidak

Pada tahapan perhitungan ini, hasil yang ingin diketahui adalah prediksi keputusan kontrak kerja dari karyawan yang memiliki NIP 264. Dimana diketahui nilai penilaian kinerjanya sebagai berikut :

- Pelayanan Prima (PP) = B
- Key Performance Indicator (KPI) = C
- Performance Appraisal (PA) = B
- Rekomendasi (R) = YA
- Keputusan = ?

Tahap 1 → Hitung Jumlah Kelas per Label
P(Ci)

$P(\text{Keputusan} = \text{Diperpanjang}) = 90 / 131$
 $P(\text{Keputusan} = \text{Tidak Diperpanjang}) = 41 / 131$

Keterangan :
 $P(\text{Keputusan} = \text{Diperpanjang})$ → jumlah data dengan Keputusan= “Diperpanjang” pada data learning dibagi dengan seluruh jumlah data learning.
 $P(\text{Keputusan} = \text{Tidak Diperpanjang})$ → jumlah data dengan Keputusan= “Tidak Diperpanjang” pada data learning dibagi dengan seluruh jumlah data learning.

Tahap 2 → Hitung jumlah kasus yang sama dengan kelas yang sama

$P(X|Ci)$
 $P(PP=B | \text{Keputusan}=\text{Diperpanjang}) = 36 / 90$
 $P(PP=B | \text{Keputusan}=\text{Tidak Diperpanjang}) = 12 / 41$
 $P(KPI=C | \text{Keputusan}=\text{Diperpanjang}) = 30 / 90$
 $P(KPI=C | \text{Keputusan}=\text{Tidak Diperpanjang}) = 7 / 41$
 $P(PA=B | \text{Keputusan}=\text{Diperpanjang}) = 19 / 90$
 $P(PA=B | \text{Keputusan}=\text{Tidak Diperpanjang}) = 13 / 41$
 $P(R=YA | \text{Keputusan}=\text{Diperpanjang}) = 66 / 90$
 $P(R=YA | \text{Keputusan}=\text{Tidak Diperpanjang}) = 31 / 41$

Keterangan :

$P(P = B | \text{Keputusan} = \text{Diperpanjang}) \rightarrow$ Jumlah data dengan G_Pelayanan Prima B dan Keputusan Diperpanjang dibagi jumlah seluruh data dengan Keputusan Diperpanjang

$P(P = B | \text{Keputusan} = \text{Tidak Diperpanjang}) \rightarrow$ Jumlah data dengan G_Pelayanan Prima B dan Keputusan Tidak Diperpanjang dibagi jumlah seluruh data dengan Keputusan Tidak Diperpanjang

$P(KPI = C | \text{Keputusan} = \text{Diperpanjang}) \rightarrow$ Jumlah data dengan KPI C dan Keputusan Diperpanjang dibagi jumlah seluruh data dengan Keputusan Diperpanjang

$P(KPI = C | \text{Keputusan} = \text{Tidak Diperpanjang}) \rightarrow$ Jumlah data dengan KPI C dan Keputusan Tidak Diperpanjang dibagi jumlah seluruh data dengan Keputusan Tidak Diperpanjang

$P(PA = B | \text{Keputusan} = \text{Diperpanjang}) \rightarrow$ Jumlah data dengan Performance Appraisal B dan Keputusan Diperpanjang dibagi jumlah seluruh data dengan Keputusan Diperpanjang

$P(PA = B | \text{Keputusan} = \text{Tidak Diperpanjang}) \rightarrow$ Jumlah data dengan Performance Appraisal B dan Keputusan Tidak Diperpanjang dibagi jumlah seluruh data dengan Keputusan Tidak Diperpanjang

$P(R = YA | \text{Keputusan} = \text{Diperpanjang}) \rightarrow$ Jumlah data dengan Rekomendasi YA dan Keputusan Diperpanjang dibagi jumlah seluruh data dengan Keputusan Diperpanjang

$P(R = YA | \text{Keputusan} = \text{Tidak Diperpanjang}) \rightarrow$ Jumlah data dengan Rekomendasi YA dan Keputusan Tidak Diperpanjang dibagi jumlah seluruh data dengan Keputusan Tidak Diperpanjang

Tahap 3 \rightarrow Kalikan semua hasil variable kelas

$$P(X|Ci) * P(Ci)$$

Untuk kelas : Diperpanjang

$$P(P=B), (KPI=C), (PA=B), (R=YA) = \{P(P=P | \text{Keputusan} = \text{Diperpanjang}), P(KPI = C | \text{Keputusan} = \text{Diperpanjang}), P(PA=B | \text{Keputusan} = \text{Diperpanjang}), P(R = YA | \text{Keputusan} = \text{Diperpanjang})\}$$

$$= (36 / 90) * (30 / 90) * (19 / 90) * (66 / 90)$$

$$= 0.02064198$$

Untuk kelas : Tidak Diperpanjang

$$P(P=B), (KPI=C), (PA=B), (R=YA) = \{P(P=P | \text{Keputusan} = \text{Tidak Diperpanjang}), P(KPI = C | \text{Keputusan} = \text{Tidak Diperpanjang}), P(PA = B | \text{Keputusan} = \text{Tidak Diperpanjang}), P(R = YA | \text{Keputusan} = \text{Tidak Diperpanjang})\}$$

$$= (12 / 41) * (7 / 41) * (13 / 41) * (31 / 41)$$

$$= 0.00184304$$

Tahap 4 \rightarrow Bandingkan hasil kelas Diperpanjang & Tidak Diperpanjang

$$P(X|Ci) * P(Ci)$$

$$P(X|\text{Keputusan} = \text{“Diperpanjang”}) * P(\text{Keputusan} = \text{“Diperpanjang”})$$

$$= 0.02064198 * 90 / 131$$

$$= 0.01396825$$

$$P(X|\text{Keputusan} = \text{“Tidak Diperpanjang”}) * P(\text{Keputusan} = \text{“Tidak Diperpanjang”})$$

$$= 0.00184304 * 41 / 131$$

$$= 0.00056816$$

Bedasarkan hasil kali semua variable kelas, hasil (P | Diperpanjang) lebih besar dari hasil (P | Tidak Diperpanjang), maka hasil prediksi yang didapat adalah “Diperpanjang”. Untuk perhitungan data testing yang lainnya dapat dihitung dengan cara yang sama seperti perhitungan diatas.

Bedasarkan tabel hasil prediksi data testing, dapat disimpulkan bahwa dari 10 data testing, karyawan yang masa kontrak kerjanya dapat Diperpanjang adalah sebanyak 9 orang karyawan, sedangkan karyawan yang masa kontrak kerjanya Tidak Diperpanjang adalah sebanyak 1 orang karyawan. Dan berdasarkan tabel data real, dapat disimpulkan bahwa dari 10 data testing, karyawan yang masa kontrak kerjanya dapat Diperpanjang adalah sebanyak 7 orang karyawan, sedangkan karyawan yang masa kontrak kerjanya Tidak Diperpanjang adalah sebanyak 3 orang karyawan.

Untuk mengetahui seberapa besar keakuratan perhitungan algoritma *naive bayes* ini digunakan, maka dibutuhkan data real atau data asli dari pengajuan perpanjangan kontrak. Berikut adalah perhitungan untuk menghitung akurasi dari hasil prediksi *naive bayes* dalam penelitian ini :

Tahap Perhitungan Akurasi

Jumlah Prediksi Diperpanjang = 9 data

Jumlah Prediksi Tidak Diperpanjang = 1 data

Jumlah Real Diperpanjang = 7 data

Jumlah Real Tidak Diperpanjang = 3 data

TP = 7 data

TN = 1 data

FP = 2 data

FN = 0 data

Keterangan :

TP = Jumlah data yang diprediksi DIPERPANJANG dan kenyataannya DIPERPANJANG

TN = Jumlah data yang diprediksi TIDAK DIPERPANJANG dan kenyataannya TIDAK DIPERPANJANG

FP = Jumlah data yang diprediksi DIPERPANJANG tapi kenyataannya TIDAK DIPERPANJANG

FN = Jumlah data yang diprediksi TIDAK DIPERPANJANG tapi kenyataannya DIPERPANJANG

$$\text{Akurasi} = (TP+TN) / (TP+TN+FP+FN) * 100\%$$

$$= (7 + 1) / (7 + 1 + 2 + 0) * 100\%$$

$$= 8/10 * 100\% = 0,8 * 100\%$$

$$= 80\%$$

$$\text{Laju error} = (FN+FP) / (TP+TN+FP+FN) * 100\%$$

$$= (0 + 2) / (7 + 1 + 2 + 0) * 100\%$$

$$= 2/10 * 100\% = 0,2 * 100\%$$

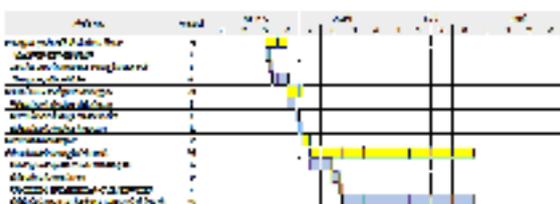
= 20%

Dengan nilai akurasi sebesar 80% dan laju error sebesar 20%, maka disimpulkan bahwa dengan prediksi 10 kasus maka hasil yang dipatkan bisa dijadikan sebagai rekomendasi untuk mengambil keputusan, sedangkan dengan adanya laju error 20% hal ini dapat disebabkan karena adaya beberapa data pengajuan perpanjangan kontrak karyawan yang mendapatkan rekomendasi TIDAK tetapi mendapatkan hasil prediksi DIPERPANJANG. Maka jika dilihat dari segi akurasi, perangkat lunak ini dapat menangani banyaknya pengajuan kontrak karyawan, dan meminimalisir waktu untuk menentukan status perpanjangan kontrak karyawan.

4. Implementasi dan Pengujian

Bab ini berisi uraian mengenai kegiatan dalam membangun/mewujudkan rancangan sistem baru secara nyata. Kegiatan yang dibahas meliputi kebutuhan sumber daya dan panduan penggunaan program.

4.1 Gantt Chart Implementasi



4.2 Lingkup dan Batasan Implementasi

Dalam sub bab ini akan menjelaskan mengenai ruang lingkup dan batasan implementasi, termasuk modul perangkat lunak yang akan diimplementasikan, fungsional perangkat lunak dan lain sebagainya termasuk lingkungan implementasi. Adapun lingkup dan batasan implementasi dalam penelitian ini antara lain :

1. Mengimplementasikan algoritma naïve bayes untuk memprediksi status kontrak kerja karyawan yang berstatus karyawan kontrak di Kospin Jasa
2. Data training yang digunakan adalah data pengajuan perpanjangan kontrak kerja dimana tanggal akhir kontrak kerja karyawan dalam 3 tahun kebelakang.
3. Data testing yang digunakan adalah data pengajuan perpanjangan kontrak kerja dimana tanggal akhir kontrak kerja karyawan dalam tahun 2018
4. Fungsi yang akan di implementasikan antara lain fungsi *login*, *import data*, proses *mining data*, *backup* hasil prediksi dan tampil hasil prediksi.
5. Dalam implementasi perangkat lunak, penulis mengimplementasikan perangkat lunak dengan berbasis desktop menggunakan bahasa pemrograman Java.

6. Aplikasi yang digunakan untuk membangun perangkat lunak menggunakan aplikasi Netbeans, Microsoft Excel, dan PHPMyAdmin.

4.3 Lingkup dan Lingkungan

Sub bab ini akan menjelaskan ruang lingkup atau batasan dari pengujian yang dilakukan dalam penelitian, dalam penelitian ini penulis akan melakukan pengujian pada fungsi-fungsi utama dari perangkat lunak yang dibuat, berikut adalah lingkup dan Batasan dari kegiatan pengujian yang dilakukan:

1. Lingkungan pengujian, prototype perangkat lunak diuji oleh pembimbing dan wakil pimpinan Kospin Jasa Bandung.
2. Fungsi yang diuji adalah modul :
 - a. Modul untuk import data
 - b. Modul untuk mengelola tahapan KDD
 - c. Modul untuk menghitung algoritma *naïve bayes*
 - d. Modul untuk menampilkan laporan berdasarkan kriteria

4.4 Implementasi Antar muka

4.3.1 Form Login



Gambar 4 Tampilan Form Login

4.3.2 Form Menu Utama



Gambar 5 Tampilan Form Menu Utama

4.3.3 Form Import Data Testing



Gambar 6 Form Import Data Setelah Data di import dan di proses

5.1 Kesimpulan

Berdasarkan analisa dan implementasi yang telah dilakukan, maka dapat ditarik kesimpulan bahwa dengan adanya perangkat lunak sistem prediksi status kontrak kerja karyawan menggunakan algoritma *naïve bayes* ini, dapat meminimalisir waktu yang dibutuhkan untuk proses penentuan status perpanjangan kontrak karyawan dan telah sesuai dengan prosedur yang ada.

5.2 Saran

Beberapa saran yang mungkin dapat dijadikan sebagai bahan pertimbangan untuk pengembangan selanjutnya disampaikan berikut ini:

1. Menerapkan algoritma klasifikasi lainnya yang memiliki keakuratan lebih baik untuk kasus ini.
2. Mengembangkan cara lain dalam pengiriman laporan hasil prediksi untuk pimpinan cabang, seperti penggunaan email, dll.

DAFTAR PUSTAKA

- [1]. Abidah, S. (2016). Analisis Komparasi Metode Tsukamoto dan Sugeno dalam Prediksi Jumlah Siswa Baru. *Jurnal Bianglala Informatika*, 4.
- [2]. Fachrizal, E., & Ginting, A. (2013). e-Jurnal Teknik Industri FT USU Vol 3, No. 1, pp. 7-14. Penilaian Kinerja Karyawan Menggunakan Metode 360 Degree Feedback Pada Hotel XYZ Medan.
- [3]. informatikalogi. (2017, April 8). Algoritma Naive Bayes. Diambil kembali dari Algoritma Naive Bayes: <https://informatikalogi.com/algoritma-naive-bayes/>
- [4]. Juliandi, A., Irfan, & Manurung, S. (2014). *Metodologi Penelitian Bisnis, Konsep dan Aplikasi*. Medan: UMSU Press.
- [5]. Kermite, R. Y., Winarno, A., & Rohmani, A. (2017). PERANCANGAN SISTEM ADMINISTRASI SEKOLAH DENGAN SMS GATEWAY BERBASIS WEB MENGGUNAKAN GAMMU PADA SMK LPI SEMARANG. *Journal of Information Systems*, 1-11.
- [6]. Liana, L. (2015). Pengujian Perangkat Lunak Menggunakan White Box Testing.
- [7]. Liantoni, F., & Nugroho, H. (2015). KLASIFIKASI DAUN HERBAL MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN KNEAREST NEIGHBOR. *Jurnal SimanteC*, 4.
- [8]. Pane, I. Z. (2014). Pengembangan Prototipe Piranti Lunak Sistem Informasi Manajemen Kegiatan Perakayasa Dengan Microsoft Excel. *ULTIMA InfoSys*, 54-60.
- [9]. Purwanto. (2017, February 6). Metodologi System Development Life Cycle (SDLC). Diambil kembali dari Medium: <https://medium.com/@purwanto.dev/metodologi-system-development-life-cycle-sdlc-2f0349df1364>
- [10]. Rismala, R., Suyanto, & Dayawati, R. N. (2017). Prediksi Data Time Series Tingkat Inflasi di Indonesia dengan Menggunakan Differential Evolution. *journals.telkomuniversity.ac.id*, 2.
- [11]. Rolly, N., & Hakiem, N. (2015). PENGEMBANGAN APLIKASI MOBILE ACADEMIC INFORMATION SYSTEM (AIS) BERBASIS ANDROID UNTUK PENGGUNA DOSEN DAN MAHASISWA. *JURNAL TEKNIK INFORMATIKA*, 2.
- [12]. Safitri, K., Waruwu, F. T., & Mesran. (2017). SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN BERPRESTASI DENGAN MENGGUNAKAN METODE ANALYTICAL HIEARARCHY PROCESS. *MEDIA INFORMATIKA BUDIDARMA*, 2.
- [13]. Saifudin, A. (2018). METODE DATA MINING UNTUK SELEKSI CALON MAHASISWA PADA PENERIMAAN MAHASISWA BARU DI UNIVERSITAS PAMULANG. *Jurnal Teknologi*, 2.
- [14]. Saleh, A. (2015). Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *Citec Journal*, 2.
- [15]. Santoso, H., Hariyadi, I. P., & Prayitno. (2016). DATA MINING ANALISA POLA PEMBELAN PRODUK DENGAN MENGGUNAKAN METODE ALGORITMA APRIORI. *Seminar Nasional Teknologi Informasi dan Multimedia 2016*, 2.
- [16]. Setiawan, S. (2017, February 13). "Studi Kepustakaan" Pengertian & (Tujuan – Peranan – Sumber – Strategi). Diambil kembali dari Guru Pendidikan: <http://www.gurupendidikan.co.id/studi->

- kepastakaan-pengertian-tujuan-peranan-sumber-strategi/
- [17]. Shanto. (2017, Mei 19). DEFINISI KONTRAK KERJA. Diambil kembali dari Serikat Pekerja Nasional (SPN): <https://spn.or.id/definisi-kontrak-kerja/>
- [18]. Suryantara, I. G. (2017). Merancang Aplikasi dengan Metodologi Extreme Programmings. Jakarta: PT Elex Media Komputindo.
- [19]. Warsito, A. B., Yusup, M., & Yulianto. (2014). KAJIAN YII FRAMEWORK DALAM PENGEMBANGAN WEBSITE PERGURUAN TINGGI. 4.