



REKOMENDASI PENSIUN DINI MENGGUNAKAN ALGORITMA NAÏVE BAYES (STUDI KASUS DI PALDAM JAYA)

Anis Zubair¹, Fadilah Ariyani²

S1 Sistem Informasi, Fakultas Teknologi Informasi, Universitas Merdeka Malang

Email: anis.zubair@unmer.ac.id¹, fadilaariyani06@gmail.com²

ABSTRAK

Penelitian ini membahas rekomendasi pensiun dini di dalam lingkungan instansi Paldam Jaya. Metode yang digunakan pada penelitian ini adalah Naïve Bayes Classifier. Data yang digunakan adalah data sampel yang terdiri atas 42 baris dan 10 kolom. Namun setelah dilakukan praproses data sampel tersebut menjadi 42 baris dan 5 kolom. Hasil penelitian menunjukkan bahwa Naïve Bayes Classifier adalah algoritma yang tepat untuk digunakan pada rekomendasi pensiun dini.

Kata kunci : Naïve Bayes Classifier, rekomendasi, pensiun dini

ABSTRACT

This study discusses the recommendations for early retirement within the Paldam Jaya agency. The method used in this study is the Naïve Bayes Classifier. The data used is sample data consisting of 42 rows and 10 columns. However, after preprocessing the sample data became 42 rows and 5 columns. The results show that the Naïve Bayes Classifier is an accurate algorithm for use in early retirement recommendations.

Keyword : Naïve Bayes Classifier, recommendation, early retireme



PENDAHULUAN

Pensiun dini atau pensiun sebelum waktunya adalah berhentinya seseorang dari bekerja untuk suatu institusi atau instansi pemerintah atau swasta karena usia atau permintaan sendiri. Bagi sebagian orang pensiun adalah hal yang sangat dinanti karena orang tersebut sudah jenuh bekerja atau ia sudah punya pekerjaan sampingan yang hasilnya jika dilakukan secara penuh waktu akan memberikan hasil yang lebih besar daripada gajinya setiap bulan. Bagi sebagian yang lain pensiun adalah hal yang ditakuti karena berkurangnya penghasilan [1]. Menurut Undang-undang Republik Indonesia Nomor 4 Tahun 2014 Tentang Aparatur Sipil Negara ada tiga batas usia pensiun. Masing-masing adalah 58 (lima puluh delapan) tahun bagi Pejabat Administrasi, 60 (enam puluh) tahun bagi Pejabat Pimpinan Tinggi, atau sesuai dengan ketentuan peraturan perundang-undangan bagi Pejabat Fungsional [2]. Data yang digunakan oleh peneliti adalah data dari Instansi Paldam Jaya. Data yang digunakan adalah data sampel. Sesuai dengan kaidah penelitian pada data mining, data akan melalui praproses sebelum dianalisis menggunakan Naïve Bayes Classifier.

Naïve Bayes Classifier adalah algoritma atau classifier yang berdasarkan teorema Bayes. Naïve Bayes Classifier dipilih dalam penelitian ini karena algoritma ini termasuk salah satu algoritma terbaik dalam klasifikasi [5]. Teorema Bayes dinyatakan dalam bentuk perbandingan peluang. Teorema tersebut dinyatakan dengan rumus di bawah ini.

$$P(A|B) = P(B|A) P(A)/P(B) \quad (1)$$

$P(A|B)$: posterior probability

$P(B|A)$: likelihood probability

$P(A)$: prior probability

$P(B)$: evidence probability [3]

Hasil klasifikasi menggunakan Naïve Bayes Classifier dinyatakan dalam bentuk confusion table. Confusion tablenya dinyatakan dalam tabel di bawah ini.

Tabel 1. Confusion Table

		Predicted Condition	
		RP	RN
Actual Condition	RP	TP	FP
	RN	FN	TN

Singkatan-singkatan pada tabel dijelaskan di bawah ini.

RP: Real Positives

RN: Real Negatives

TP: True Positives

FP: False Positives



FN: False Negatives

TN: True Negatives

Recall, precision, f-measure, dan accuracy dapat dihitung dari confusion table di atas. Di bawah ini adalah rumus untuk recall, precision, f-measure, dan accuracy.

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FP})$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FN})$$

$$\text{F-measure} = 2 * \text{recall} * \text{precision}/(\text{recall} + \text{precision})$$

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{FN} + \text{TN}) [2]$$

METODE

Objek penelitian ini adalah Instansi Paldam Jaya yang terletak di Jl. Mayjen Sutoyo No.05, Cililitan, Kramat Jati, RT.7/RW.7, Cawang, Kec. Kramat Jati, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta. Data mentah yang diambil adalah data sekunder. Data sekunder terdiri atas 42 baris dan 10 kolom. Data sekunder ini merupakan sampel dari data yang ada dalam lingkungan instansi Paldam Jaya. Penjelasan data tersebut dinyatakan di tabel di bawah ini.

Tabel 2. Penjelasan Data Mentah

JUDUL	PENJELASAN
Nama	Nama lengkap
Tanggal Lahir	Tanggal dalam format dd-mm-yyyy
Jenis Kelamin	Laki-laki atau perempuan
NIP	NIP dalam bentuk 12 angka
Golongan	Golongan dalam bentuk angka romawi/huruf
Jabatan	Jabatan yang ada di Paldam Jaya
Satuan	Paldam Jaya
Pendidikan	SMA atau setara sampai dengan S1 atau setara
Tanggal Masuk	Tanggal dalam format dd-mm-yyyy
Pensiun Dini	Ya atau tidak

Selanjutnya dilakukan praproses data. Kolom Nama, NIP, Jabatan, dan Satuan dihilangkan. Kolom Tanggal Lahir diubah menjadi Umur dalam bentuk rentang. Pengubahan ini menggunakan rumus Sturges. Kolom Tanggal Masuk diubah menjadi Masa Kerja. Hasil praproses data terdiri atas 42 baris dan 5 kolom. Penjelasan data tersebut dinyatakan di tabel di bawah ini.

Tabel 3. Penjelasan Data Hasil Praproses

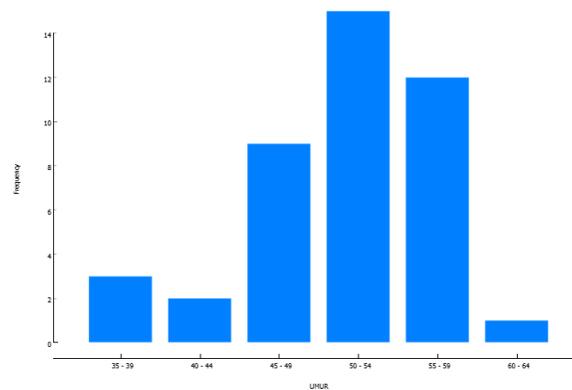
JUDUL	PENJELASAN
Umur	Umur dalam bentuk rentang
Jenis Kelamin	Laki-laki atau perempuan
Golongan	Golongan dalam bentuk angka romawi/huruf
Pendidikan	SMA atau setara sampai dengan S1 atau setara
Masa Kerja	Lebih dari 20 tahun atau tidak
Pensiun Dini	Ya atau tidak



Metode yang digunakan dalam penelitian ini adalah Naïve Bayes. Langkah pertama algoritma Naïve Bayes adalah menghitung likelihood probability. Selanjutnya adalah menghitung prior probability. Langkah berikutnya adalah menghitung posterior probability. Setelah menghitung posterior probability langkah selanjutnya adalah memasukkan hasilnya ke dalam bentuk confusion table. Langkah terakhir adalah menghitung recall, precision, f-measure, dan accuracy dari confusion table tersebut.

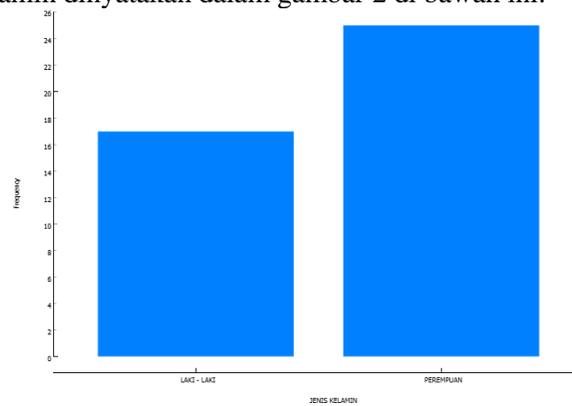
HASIL DAN PEMBAHASAN

Deskripsi visual data ini dinyatakan dalam bentuk diagram batang. Deskripsi visual Umur dinyatakan dalam gambar 1 di bawah ini.



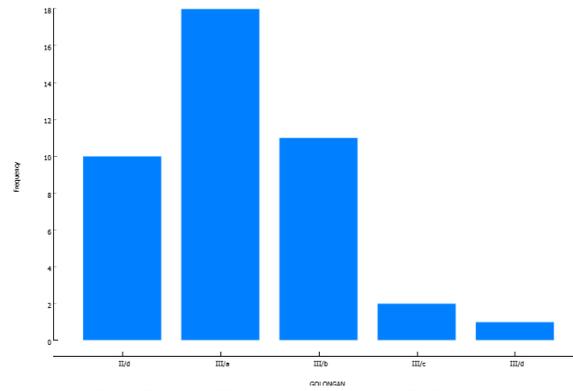
Gambar 1. Diagram Batang Umur

Deskripsi visual Jenis Kelamin dinyatakan dalam gambar 2 di bawah ini.



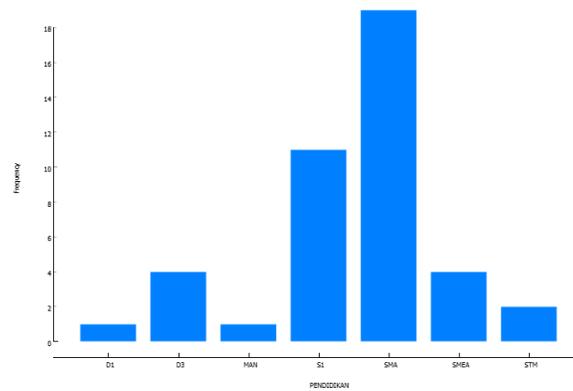
Gambar 2. Diagram Batang Jenis Kelamin

Deskripsi visual Golongan dinyatakan dalam gambar 3 di bawah ini.



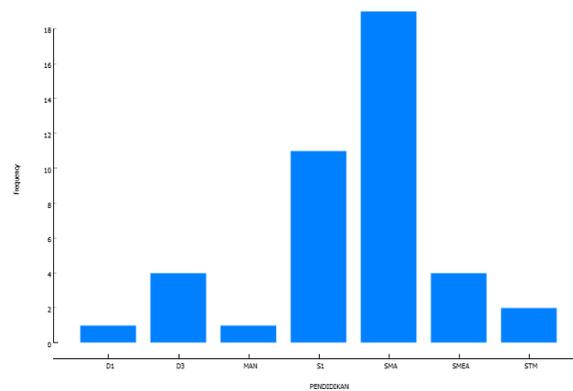
Gambar 3. Diagram Batang Golongan

Deskripsi visual Pendidikan dinyatakan dalam gambar 4 di bawah ini.



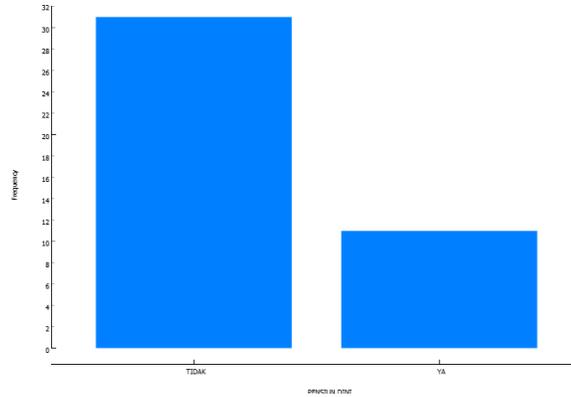
Gambar 4. Diagram Batang Pendidikan

Deskripsi visual Masa Kerja dinyatakan dalam gambar 5 di bawah ini.



Gambar 5. Diagram Batang Masa Kerja

Deskripsi visual Pensiun Dini dinyatakan dalam gambar 6 di bawah ini.



Gambar 6. Diagram Batang Pensiun Dini

Langkah pertama algoritma Naïve Bayes adalah menghitung likelihood probability. Likelihood probability Umur dinyatakan dalam tabel di bawah ini.

Tabel 3. Likelihood Probability Umur

UMUR	TIDAK	YA
55 - 59	0,387097	0
60 - 64	0,032258	0
50 - 54	0,419355	0,181818
35 - 39	0,096774	0
40 - 44	0,064516	0
45 - 49	0	0,818182

Likelihood probability Jenis Kelamin dinyatakan dalam tabel di bawah ini.

Tabel 4. Likelihood Probability Jenis Kelamin

JENIS KELAMIN	TIDAK	YA
LAKI - LAKI	0,483871	0,181818
PEREMPUAN	0,516129	0,818182

Likelihood probability Golongan dinyatakan dalam tabel di bawah ini.

Tabel 5. Likelihood Probability Golongan

GOLONGAN	TIDAK	YA
III/d	0,032258	0
III/c	0,064516	0
III/b	0,354839	0
III/a	0,387097	0,545455
II/d	0,16129	0,454545

Likelihood probability Pendidikan dinyatakan dalam tabel di bawah ini.

Tabel 6. Likelihood Probability Pendidikan

PENDIDIKAN	TIDAK	YA
S1	0,064516	0
D3	0,129032	0
S1	0,258065	0,090909



SMA	0,322581	0,636364
STM	0,064516	0
D1	0,032258	0
MAN	0,032258	0
SMEA	0,032258	0,090909
SMA	0,064516	0
SMEA	0	0,181818

Likelihood probability Masa Kerja dinyatakan dalam tabel di bawah ini.

Tabel 7. Likelihood Probability Masa Kerja

MASA KERJA	TIDAK	YA
MK > 20	0,870968	1
MK < 20	0,129032	0

Langkah kedua adalah menghitung prior probability. Prior probability Pensiun Dini dinyatakan dalam tabel di bawah ini.

Tabel 8. Prior Probability Pensiun Dini

TIDAK	YA
0,738095	0,261905

Dengan menggunakan rumus (1) maka posterior probability dinyatakan dalam tabel di bawah ini.

Tabel 9. Posterior Probability

NOMOR	TIDAK	YA
1	0,00034	0
2	0,000113	0
3	0,002716	0
4	0,019918	0
5	0,019918	0
6	0,003735	0
7	0,018673	0
8	0,021578	0
9	0,021578	0
10	0,018673	0
11	0,007967	0
12	0,020229	0
13	0,017262	0
14	0,000553	0
15	0,002173	0
16	0,004414	0
17	0,000604	0
18	0,000604	0
19	0,020371	0
20	0,002354	0
21	0,002897	0



NOMOR	TIDAK	YA
22	0,021729	0
23	0,000215	0
24	0,004708	0
25	0,018832	0,007377
26	0,017655	0,001639
27	0,003621	0
28	0,000981	0,006147
29	0,001839	0
30	0,009195	0,009562
31	0,001962	0
32	0	0,051636
33	0	0,232361
34	0	0,232361
35	0	0,033194
36	0,022068	0,011475
37	0	0,232361
38	0	0,193634
39	0	0,193634
40	0	0,055324
41	0	0,055324
42	0,000981	0,006147

Langkah berikutnya adalah membandingkan posterior probability TIDAK dan YA. Posterior probability yang dipilih adalah yang lebih besar. Tabel di bawah ini menunjukkan prediksi pensiun dini atau tidak.

Tabel 10. Prediksi Pensiun Dini atau Tidak

NOMOR	PENSIUN DINI	PREDIKSI
1	TIDAK	TIDAK
2	TIDAK	TIDAK
3	TIDAK	TIDAK
4	TIDAK	TIDAK
5	TIDAK	TIDAK
6	TIDAK	TIDAK
7	TIDAK	TIDAK
8	TIDAK	TIDAK
9	TIDAK	TIDAK
10	TIDAK	TIDAK
11	TIDAK	TIDAK
12	TIDAK	TIDAK
13	TIDAK	TIDAK
14	TIDAK	TIDAK
15	TIDAK	TIDAK
16	TIDAK	TIDAK
17	TIDAK	TIDAK



NOMOR	PENSIUN DINI	PREDIKSI
18	TIDAK	TIDAK
19	TIDAK	TIDAK
20	TIDAK	TIDAK
21	TIDAK	TIDAK
22	TIDAK	TIDAK
23	TIDAK	TIDAK
24	TIDAK	TIDAK
25	TIDAK	TIDAK
26	TIDAK	TIDAK
27	TIDAK	TIDAK
28	TIDAK	YA
29	TIDAK	TIDAK
30	TIDAK	YA
31	TIDAK	TIDAK
32	YA	YA
33	YA	YA
34	YA	YA
35	YA	YA
36	YA	TIDAK
37	YA	YA
38	YA	YA
39	YA	YA
40	YA	YA
41	YA	YA
42	YA	YA

Setelah menunjukkan prediksi pensiun dini atau tidak langkah selanjutnya adalah memasukkan hasilnya ke dalam bentuk confusion table.

Tabel 11. Confusion Table

		Prediksi	
		TIDAK	YA
Kenyataan	TIDAK	29	2
	YA	1	10

Langkah terakhir adalah menghitung recall, precision, f-measure, dan accuracy dari confusion table tersebut. Hasilnya ditunjukkan oleh tabel di bawah ini.

Tabel 12. Recall, Precision, F-Measure, dan Accuracy

Recall	94%
Precision	97%
Accuracy	93%
F-measure	95%



SIMPULAN

Berdasarkan hasil penelitian ini klasifikasi menggunakan Algoritma Naïve Bayes dapat digunakan untuk membuat rekomendasi pensiun dini pada pegawai yang ada dalam lingkungan instansi Paldam Jaya. Hasil recall, precision, f-measure, dan accuracy klasifikasi menggunakan Algoritma Naïve Bayes termasuk akurat karena di atas 90%. Penelitian ini menggunakan data sampel yang terdiri atas 42 baris. Saran yang dapat dijadikan pertimbangan untuk penelitian berikutnya adalah menggunakan semua data yang ada dalam lingkungan instansi Paldam Jaya.

DAFTAR RUJUKAN

- 7 Langkah Merencanakan Pensiun Dini dengan Baik. <https://www.jojonomic.com/blog/pensiun-dini/#:~:text=Ketentuan%20Pensiun%20Dini,karena%20usia%20atau%20permintaan%20sendiri>.
- Powers, D. M. W. (2011). Evaluation: from Precision, Recall, and F-measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*. 2(1): 37-63.
- Tan, P. N., Steinbach, M., & Kumar, V. (2013). *Introduction to Data Mining*. Harlow: Pearson Education Limited.
- Undang-undang Republik Indonesia Nomor 4 Tahun 2014 Tentang Aparatur Sipil Negara. <https://jdih.kemenkeu.go.id/fulltext/2014/5TAHUN2014UU.htm>.
- Zhang, C., et. Al. (2017). An up-to-date Comparison of State-of-the-art Classification Algorithms. *Expert Systems with Applications* 82: 128-150