

Klasifikasi Level Stress Manusia menggunakan Algoritma Naïve Bayes

Taufik Abdul Rahman^{1*}, Al Habib Galih Siswoyo², Paduloh³

^{1,2,3}Program Studi Teknik Industri, Fakultas Teknik, Universitas Bhayangkara Jakarta Raya, Jakarta, Indonesia

e-mail: *¹taufik.abdul.rahman19@mhs.ubharajaya.ac.id,

²al.habib.galih.siswoyo19@mhs.ubharajaya.ac.id, ³paduloh@dsn.ubharajaya.ac.id

Abstract

Companies in carrying out its activities require employees who are healthy both physically and mentally therefore employees as HR is the most important part for the company, when employees experience stress will be bad for the productivity of the company in this study aims to classify the level of stress on employees. This study uses classification techniques and stages in data mining for data classification of stress levels in employees with the algorithm Naive Bayes using RapidMiner tool, data processing will be used as a dataset in this study. Of the data will be divided into 70% training data and 30% testing data. The results of the research data states the level of performace/accuracy 97.67%, simple distribution that contains the normal class 0.395, attribute temperature in the normal class is 88.065, decide in the classification of Naive bayes. Based on the research that has been done to get the conclusion that the classification technique with the application of the algorithm Naive Bayes can be used to predict the level of stress levels in employees.

Keywords : Data Mining, Naïve Bayes, classification, RapidMiner

Abstrak

Perusahaan dalam menjalankan aktifitasnya membutuhkan karyawan yang sehat baik fisik maupun mental maka dari itu karyawan sebagai SDM merupakan bagian terpenting untuk perusahaan, ketika karyawan mengalami stress akan berakibat buruk terhadap produktifitas perusahaan dalam penelitian ini bertujuan untuk mengklasifikasi level stress terhadap karyawan. Penelitian ini menggunakan teknik klasifikasi dan tahapan - tahapan pada data mining untuk klasifikasi data level stress pada karyawan dengan algoritma Naïve Bayes menggunakan tool RapidMiner, pengolahan data yang akan dijadikan dataset dalam penelitian ini. Dari data tersebut akan dibagi menjadi 70% data training dan 30% data testing. Hasil dari penelitian data menyatakan tingkat performace / accuracy 97,67 %, simple distribution yang berisi class normal 0,395, attribute temperature pada class normal adalah 88,065, putuskan dalam klasifikasi naïve bayes. Berdasarkan penelitian yang sudah dilakukan mendapatkan kesimpulan bahwa teknik klasifikasi dengan penerapan algoritma Naïve Bayes dapat digunakan untuk melakukan prediksi tingkat level stress pada karyawan.

Kata Kunci: Data Mining, Naïve Bayes , Klasifikasi, RapidMiner

PENDAHULUAN

Persaingan dan tuntutan profesionalitas yang semakin tinggi menimbulkan banyaknya tekanan-tekanan yang harus dihadapi individu dalam lingkungan kerja. Dampak yang sangat merugikan dari adanya gangguan kecemasan yang sering dialami oleh masyarakat dan karyawan khususnya disebut stress. Menurut Wirawan (2012) stress merupakan reaksi yang tidak diharapkan muncul sebagai akibat tingginya tuntutan lingkungan kepada seseorang. Stress kerja merupakan aspek yang penting bagi perusahaan terutama keterkaitannya dengan kinerja karyawan. Menurut Mangkunegara (2009) kinerja adalah asil kerja secara kualitas dan kuantitas yang dicapai oleh seorang karyawan dalam melaksanakan tugasnya sesuai dengan tanggung jawab yang diberikan kepadanya.

Perusahaan mengalami penurunan produktifitas maka salah satu indikasi permasalahan ini adalah tingkat stress pada karyawan, peneitian ini bertujuan untuk mengklasifikasi level stress pada karyawan PT. XYZ agar dapat diketahui penyebab dari penurunan produktifitas tersebut. Peneitian ini menggunakan objek manusia lebih tepat-nya adalah karyawan di salah satu perusahaan yang

bernama PT. XYZ daerah Cikarang Kabupaten Bekasi

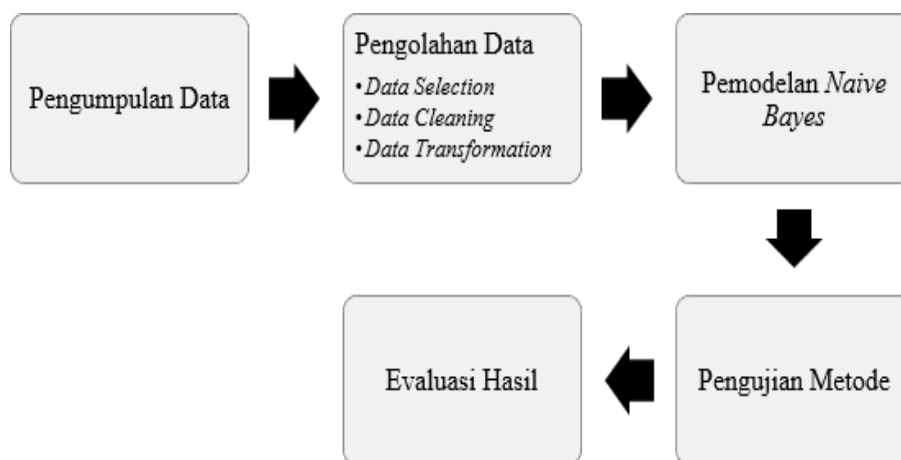
Untuk mengetahui data yang dibutuhkan untuk klasifikasi level stress Peneliti menganalisa langsung di PT. XYZ dibantu dengan beberapa pihak dalam PT. XYZ tersebut bertujuan untuk mempermudah mengambil data yang akan digunakan untuk proses klasifikasi data level stress terhadap karyawan, atribut yang digunakan dalam data ini adalah kelembapan, temperatur dan perilaku tidak jelas.

Algoritma naïve bayes merupakan salah satu algoritma yang digunakan pada *text Mining* atau *data mining* (Darwis et al., 2021). Algoritma *Naïve Bayes* adalah algoritma yang digunakan dalam mengklasifikasikan data (Mariyanto et al., 2022). Kelebihan dari algoritma *naïve Bayes* adalah dapat memprediksi probabilitas beberapa kelas variabel target (Rachman & Handayani, 2021). Algoritma *Naïve Bayes* hanya membutuhkan jumlah data training yang sedikit dalam menentukan estimasi parameter pada proses klasifikasinya (Nurmayanti et al., 2021). Naive Bayes digunakan untuk teknik pengklasifikasian data dengan metode kemungkinan dan statistik yang memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive Bayes dimana diasumsikan kondisi antara atribut yang satu dengan atribut yang lainnya itu saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak adanya ciri-ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas yang lainnya (Mariyanto et al., 2022), (Silvana et al., 2020), (Syaputri et al., 2020), (Asfi & Fitrianiingsih, 2020).

Pada penelitian ini metode yang digunakan adalah Naïve Bayes untuk melakukan klasifikasi level stress karyawan pada PT. XYZ

METODE PENELITIAN

Dalam melakukan analisis dan mencari pola pada data karyawan yang mengalami stress agar memudahkan penelitian dan dapat berjalan dengan sistematis dan memenuhi tujuan yang diinginkan maka dibuat langkah – langkah dalam tahapan penelitian yang akan dilakukan berikut :



Gambar 1. Tahapan Penelitian

Pada penelitian ini menggunakan teknik klasifikasi dan tahapan - tahapan pada data mining untuk klasifikasi data level stress dengan algoritma Naïve Bayes yang terdiri dari :

1. Pengolahan data dan yang akan dijadikan dataset dalam penelitian ini. Dari data tersebut akan dibagi menjadi 70% data *training* dan 30% data *testing*. Memisahkan data menjadi training dan *testing set* dimaksudkan agar model yang diperoleh nantinya memiliki kemampuan generalisasi yang baik dalam melakukan klasifikasi data. Data training atau *training set* adalah bagian dataset yang dilatih untuk membuat klasifikasi atau menjalankan fungsi dari sebuah algoritma sesuai dengan tujuannya masing-masing. Data *testing* atau *test set* adalah bagian dataset yang digunakan untuk melihat keakuratan atau performa dalam hasil.
2. Pengolahan data, terdapat 3 tahapan yang terdiri dari :
 - a. *Data Selection*, merupakan proses pemilihan data dari sekumpulan data operasional yang ada sebelum masuk ke tahap mining data maupun informasi. Pada proses ini dilakukan pemilihan data atribut atau variable yang relevan yang akan digunakan dalam penelitian.

- Karena tidak semua atribut yang terdapat dalam database dapat digunakan. untuk melakukan penelitian yaitu dengan memilih atribut – atribut yang akan digunakan dan menghilangkan atribut-atribut yang kurang relevan atau terdapat data yang rusak.
- b. *Data cleaning*, Pada tahap data *cleaning* merupakan proses pembersihan dari data yang akan dipakai untuk penghapusan data dengan membuang *missing value*, duplikasi data, dan memeriksa inkonsistensi data dan memperbaiki kesalahan pada data. Proses pembersihan data dilakukan secara manual dengan bantuan *software spreadsheet*.
 - c. *Data transformation*, Tahap Data Transformation merupakan proses mengubah format data awal menjadi sebuah format data standar untuk proses pembacaan data dengan algoritma pada program maupun *tool* yang digunakan. dalam pengolahan dan dapat dihasilkannya hasil akurasi yang baik maka data harus di *transforming* kedalam bentuk data yang mudah di pahami. Pada penelitian ini data hasil unduhan diubah kedalam data kualitatif yang dapat mempermudah dalam pemodelan.
3. Pemodelan *Naive Bayes*, Pemodelan pada penelitian ini dilakukan dengan data mining teknik klasifikasi algoritma *Naive Bayes*. Teknik ini dipilih karena merupakan metode yang umum dipakai pada penelitian data mining untuk mengklasifikasi atau mengenali data - data yang dipelajari terutama pada prediksi tingkat level stres manusia. Algoritma yang akan diterapkan pada penelitian ini adalah *Naive Bayes*. Algoritma ini merupakan algoritma yang sudah mapan dan banyak diimplementasikan pada teknik klasifikasi. Selain itu algoritma ini memiliki kelebihan yaitu berupa akurasi yang baik dalam menangani sebuah dataset yang diolah.
 4. Pengujian metode dilakukan dengan maksud mengetahui hasil perhitungan yang dianalisa dan mengukur metode serta algoritma yang digunakan apakah berfungsi dengan baik atau tidak. Proses pengujian menggunakan tool rapidminer dan melihat data apakah sesuai dengan hasil yang diperoleh melalui *tool* tersebut. Sedangkan validasi metode dan algoritma *Naive Bayes* dilakukan dengan mengukur hasil *accuracy*.

HASIL DAN PEMBAHASAN

Penelitian ini adalah pengujian terhadap algoritma *naive bayes*, untuk mengetahui algoritma yang mana yang akan mendapatkan hasil nilai *accuracy*, *simple distribution*, *distribution table* yang dapat digunakan untuk mengetahui level stress karyawan. Data yang digunakan dalam penelitian ini terdiri dari atribut atau variabel seperti: Kelembapan, Temperatur dan perilaku tidak jelas.

Data yang digunakan dalam penelitian ini merupakan data Sekunder dan kualitatif yang didapatkan dari PT. XYZ yang terletak di Cikarang Kabupaten Bekasi.

Tabel 1 Data Stress Karyawan

<i>Humidity</i>	<i>Temperature</i>	<i>Step count</i>	<i>Stress Level</i>
21.33	90.33	123	1
21.41	90.41	93	1
27.12	96.12	196	2
27.64	96.64	177	2
10.87	79.87	87	0
11.31	80.31	40	0
18.16	87.16	88	1
28.2	97.2	162	2
14.25	83.25	61	0
26.13	95.13	168	2
23.61	92.61	200	2
19.37	88.37	117	1
29.08	98.08	179	2
17.83	86.83	55	1
28.06	97.06	148	2
19.43	88.43	123	1
26.85	95.85	169	2

26.51	95.51	135	2
29.49	98.49	149	2

Setelah data stress dimasukan, proses data seperti (subproses, *naïve bayes*, *apply model*, dan *performace*) telah dilakukan begitu juga dengan parameter sudah digunakan Menggunakan *software* Rapid miner didapatkan hasil :

Row No.	Stress Level	prediction(Stress Level)	confidence...	confidence...	confidence...	Temperature	Step count
1	normal	normal	0.973	0.927	0.900	90.330	123
2	high	high	0.906	1.000	0	97.200	162
3	normal	normal	0.999	0.901	0.900	90.430	123
4	high	high	0.930	1.000	0	90.490	149
5	normal	normal	0.988	0.912	0.900	90.540	122
6	high	high	0.930	1.000	0	97.660	131
7	normal	normal	1.000	0.900	0.900	97.620	43
8	low	low	0.252	0.900	0.917	90.430	79
9	high	high	0.930	1.000	0	95.690	152
10	normal	normal	0.902	0.900	0.430	84.870	21
11	normal	normal	0.991	0.900	0.909	86.560	30
12	high	high	0.912	0.989	0	92.540	162
13	normal	normal	1.000	0.900	0.900	88.060	87
14	high	high	0.931	0.999	0	94.460	172

Gambar 2. Data tes / Data Uji

Data tes yang digunakan 30 % terdapat 600 *exemple*.

Simple Distribution
Distribution model for label attribute Stress Level

- Class normal (0.395)
 - 2 distributions
- Class high (0.355)
 - 2 distributions
- Class low (0.251)
 - 2 distributions

Gambar 3. *Simple Distribution*

Terdapat *simple distribution* yang berisi :

- Class normal 0,395
- Class high 0,355
- Class low 0,251

Dapat diketahui bahwa class normal (tingkat stress) sebagai klasifikasi tertinggi.

Attribute	Parameter	normal	high	low
Temperature	mean	88,065	95,520	91,473
Temperature	standard deviation	2,311	2,026	1,464
Step count	mean	30,125	169,127	42,014
Step count	standard deviation	37,129	26,719	28,377

Gambar 4. Distribution Table

Terdapat *distribution table* untuk mengetahui parameter standar deviasi dan rata - rata dari *class normal, high* dan *low* yang ada pada atribut.

Kita bisa ketahui *attribute temperature* pada *class normal* adalah 88,065 dengan standar deviasinya yaitu 2,311 berarti hasil rata - rata sudah melewati standar deviasi *temperature*.

	true normal	true high	true low	class precision
pred normal	225	0	2	90.92%
pred high	0	213	0	100.00%
pred low	12	0	140	92.50%
class recall	94.94%	100.00%	93.87%	

Gambar 5. Performance

Terdapat *performance*, bisa diketahui akurasi data yang dihasilkan adalah 97,67 % merupakan tingkat akurat data yang di peroleh cukup baik.

Penelitian sebelumnya menyatakan bahwa metode *naive bayes* mempunyai hasil dengan tingkat keakuratan yang tinggi sehingga dapat menyelesaikan masalah. Hasil penelitian ini mendukung penelitian yang dilakukan Chandra (2012) dalam penelitiannya menjelaskan bahwa stres kerja secara parsial mempunyai pengaruh yang negatif dan signifikan terhadap kinerja karyawan di PT. Lie Fung Surabaya, artinya ketika stres kerja rendah maka kinerja tinggi demikian sebaliknya. Alishah (2015) dalam penelitiannya menunjukkan adanya pengaruh yang signifikan antara variabel stres kerja dan kinerja karyawan di Koperasi Syari'ah Binama Semarang.

KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian dengan menggunakan algoritma *naive bayes* dapat diambil suatu kesimpulan sebagai berikut :

Teknik klasifikasi dengan penerapan algoritma *Naive Bayes* dapat digunakan untuk melakukan prediksi tingkat level stress pada karyawan. Hasil dari penelitian data menyatakan tingkat tingkat *performace / accurcy* 97,67 %, *simple distribution* yang berisi *class normal* 0,395, *attribute temperature* pada *class normal* adalah 88,065, diputuskan dalam klasifikasi *naive bayes*

DAFTAR PUSTAKA

Asfi, M., & Fitrianiingsih, N. (2020). Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi. *InfoTekJar Jurnal Nasional Informatika Dan Teknologi Jaringan*, 5(1), 44.

Aslihah, N, Pengaruh Stres Kerja Terhadap Kinerja Karyawan Di Koperasi Syari'ah Binama Semarang.

Semarang: Universitas Islam Negeri Walisongo, 2015.

- A. Saleh, "Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga," *Creat. Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2015.
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131. <https://doi.org/10.33365/jtk.v15i1.744>
- Mariyanto, J., Pramadi, S., Ikhsan, K., Yose, I., Raywir, E., Marrylinteri, Istoningtyas, & Pratama, Y. (2022). Implementasi Algoritma Naive Bayes Untuk Memprediksi zona dan Tingkat Penyebaran Covid-19 Di Provinsi Jambi. *Jurnal Informatika Dan Rekayasa Komputer*, 2(2), 248–253.
- Nurmayanti, W. P., Saky, D. A. L., Malthuf, M., Gazali, M., & Hirzi, R. H. (2021). Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak. *Geodika: Jurnal Kajian Ilmu Dan Pendidikan Geografi*, 5(1), 123–132. <https://doi.org/10.29408/geodika.v5i1.3430>
- Rachman, R., & Handayani, R. N. (2021). Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM. *Jurnal Informatika*, 8(2), 111–122. <https://doi.org/10.31294/ji.v8i2.10494>
- Silvana, M., Akbar, R., & Syahnum, A. (2020). Pemanfaatan Metode Naïve Bayes dalam Implementasi Sistem Pakar Untuk Menganalisis Gangguan Perkembangan Anak. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(2), 74–81. <https://doi.org/10.25077/teknosi.v6i2.2020.74-81>
- Syaputri, A. W., Irwandi, E., & Mustakim, M. (2020). Naïve Bayes Algorithm for Classification of Student Major's Specialization. *Journal of Intelligent Computing & Health Informatics*, 1(1), 17. <https://doi.org/10.26714/jichi.v1i1.5570>